77

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, C07K 14/315, 19/00, C12N 15/70, 1/21, A61K 39/09

A1

(11) International Publication Number: **WO 99/35270**

(43) International Publication Date: 15 July 1999 (15.07.99)

(21) International Application Number: PCT/CA98/01203

(22) International Filing Date: 29 December 1998 (29.12.98)

(30) Priority Data:
09/001,737 31 December 1997 (31.12.97) US

(71) Applicant *(for all designated States except US)*: STRESSGEN BIOTECHNOLOGIES CORPORATION [CA/CA]; # 120 – 4243 Glanford Avenue, Victoria, British Columbia V8Z 4B9 (CA).

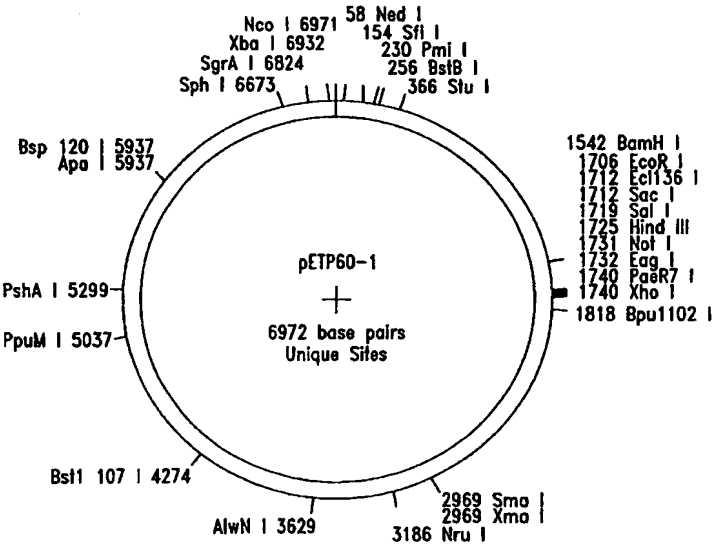
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MIZZEN, Lee [CA/CA]; 1936 Quamichan Street, Victoria, British Columbia V8S 2C4 (CA). WISNIEWSKI, Jan [PL/CA]; 5004 Tara Place RR2, Sooke, British Columbia V0S 1N0 (CA).

(74) Agents: NASSIF, Omar, A. et al.; Gowling, Strathy & Henderson, Suite 4900, Commerce Court West, Toronto, Ontario M5L 1J3 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: STREPTOCOCCAL HEAT SHOCK PROTEINS OF THE HSP60 FAMILY



(57) Abstract

Methods and compositions comprising isolated nucleic acid molecules specific to *Streptococcus pneumoniae* and *Streptococcus pyogenes*, as well as vector constructs and isolated polypeptides specific to *Streptococcus pneumoniae* and *Streptococcus pyogenes* are provided. Such compositions and methods are useful for the diagnosis of Streptococcal infection and for generating an immune response to Streptococcal bacteria.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

STREPTOCOCCAL HEAT SHOCK PROTEINS OF THE HSP60 FAMILY

TECHNICAL FIELD OF THE INVENTION

This invention relates to Streptococcal Hsp60 proteins, including fragments thereof, and nucleic acid molecules encoding such proteins and fragments, in particular from *Streptococcus pneumoniae* and *Streptococcus pyogenes,* and uses of such proteins and nucleic acid molecules.

BACKGROUND OF THE INVENTION

The World Health Organization has estimated that, worldwide, about 30% of deaths of children under age 5, or about 4-5 million, result from acute respiratory infections. David Klein, *Pneumococcal Conjugate Vaccines: Review and Update, in Microbial Drug Resistance 1*:49, 1995. The most frequent causative agent responsible for these deaths is *Streptococcus pneumoniae,* which is also referred to as pneumococcus and causes a wide variety of infections such as sinusitis, otitis, pneumonia, bacteremia and meningitis. This organism is found on respiratory mucosal membranes of 15-35% of healthy children and up to 80% of children with respiratory infections. Gray et al., *J. Infect. Dis. 142*:923, 1980; Hendley et al., *J. Infect. Dis. 132*:55, 1975. In addition, *Streptococcus pneumoniae* is responsible for 70,000 meningitis deaths and a similar number of deaths from sepsis and other infections.

In developing countries, Pneumococcal infections are responsible for approximately 1.2 million deaths among children 5 years of age and younger, which corresponds to nearly 40% of all pneumonia related deaths. World Health Organization, *Pneumococcal Conjugate Vaccines, reported in* Report of Meeting of November 15-17, 1993 (WHO/ARI/94.34). In the industrialized world, taking the U.S. as an example, pneumococcus is a leading cause of severe morbidity in the general population and of death in the elderly as well as the immunocompromised population. Klein et al. (*supra*). Pneumococcus causes more deaths (about 50,000) in older adults than any other infectious agent. High risk individuals include those with sickle cell anemia, nephrotic syndrome, asplenia, alcoholism and HIV infection. Pneumococcus

also poses a large risk to children under the age of two. In infants below the age of two, pneumococcus is the predominant cause of meningitis, bacteremia and otitis media. Within the first two years of life, about 25% of children experience otitis media caused by pneumococcus, a percentage that increases to 75% by the age of six. In Finland, two-year old children have experienced, on average, more than one episode of otitis media. About half of the cases of acute otitis media were determined to be caused by pneumococcus. Eskola, J. and Kaeyhty, H., *Ann. Med. 27*:53, 1995.

Pneumococcus is a gram-positive organism that has type-specific capsular polysaccharides. Eighty-three different type specificities have been identified and have been designated 1-83 in the American system. Jennings, *Current Topics in Microbiology and Immunology 150*:97-121, 1990. The structures of the different Pneumococcal polysaccharides have been reviewed by Kenne and Lindberg in *The Polysaccharides 2*:282-363 (Aspinal ed., 1983).

The need for an effective way to generate an immune response against *Streptococcus pneumoniae* was recognized long ago. In 1945 it was demonstrated that isolated capsular polysaccharides were able to provide type-specific protection in humans. MacLeod et al., *J. Exp. Med. 82*:445-65, 1945. However, this protection was inadequate due to the large number of different polysaccharides needed for complete protection. The interest in a vaccine soon subsided due to the success of antibiotic treatment of infections.

Recently the interest in the development of an effective vaccine has renewed. One reason was that antibiotic treatment of infectious diseases caused by encapsulated bacteria such as pneumococcus did not always prevent morbidity and mortality. For an analogous example, cured *Hemophilus influenzae* meningitis was a major cause of acquired mental retardation. Sell et al., *Pediatrics 49*:206-11, 1972. Another important reason for the renewed interest in vaccine development was the appearance and rapid spread of antibiotic-resistant strains of pneumococcus. For example, in two hospitals in Paris, France, the frequency of resistant isolates from patients increased from 1.8% in 1987 to 17% in 1990. In Barcelona, Spain, the rate of resistance increased from 4.3% in 1979 to 40% in 1990. *See* Lonks and Medeiros,

Antimicrobial Therapy 1 79:523-35, 1995. Multidrug-resistant pneumococcus have also appeared in many countries including 18 of the 50 states of the United States.

A vaccine containing polysaccharide antigens for 14 of the 83 capsular types was developed and released in 1978. Lonks and Medeiros, *supra*. This vaccine, was improved in 1983 by the creation of second generation vaccine containing 23 different polysaccharides. However, two large studies, using this vaccine, one with 2837 patients, showed that the improved vaccine was only about 57% efficacious against Pneumococcal bacteremia. Butler et al., *JAMA 270*:1826, 1993.

A drawback to polysaccharide-based vaccines is that the efficacy of these vaccines is problematic in infants under two years of age, who respond very poorly to these vaccines. Gotschlich et al., *Antibodies in Human Diagnosis and Therapy* 391-402 (Haber and Krause eds., 1977). An additional drawback is that antibodies produced by polysaccharide-based vaccines are predominantly of the IgM isotype, and therefore the immune response is not heightened upon secondary exposure to the antigen.

These and other concerns about polysaccharide-based vaccines demonstrate that there is a need in the art for improved compositions which can be used to generate an immunogenic response directed to *Streptococcus pneumoniae*.

Turning to *Streptococcus pyogenes*, also referred to as group A *streptococcus* ("GAS"), it too is a gram-positive bacterium that is causatively associated with a number of human disease states, ranging from acute pharyngitis (strep throat) to invasive diseases involving degeneration of the heart valves (acute rheumatic fever) and acute post-Streptococcal glomerulonephritis. Facklam, *Development of Group A* Streptococcal *Vaccines, in Manual of Clinical Microbiology* 1-22 (Lennette, Balows, Hausler and Truant eds., 1980). Infection by this bacterium can also cause impetigo (a supporative mucosal infection), invasive fasciitis (*viz.* flesh-eating disease), boils and skin abscesses (pyoderma), scarlet fever, sepsis, a severe toxic-shock like syndrome and pneumonia.

Before the advent of antibiotic therapy, rheumatic fever was a leading cause of mortality in children and of chronic heart disease in individuals who survived

systemic infection. In developing countries, rheumatic fever is still an enormous problem. It has been estimated that in India over 6 million school age children suffer from rheumatic heart disease. Agarwal, *Lancet I* 910-11, 1981. In the United States, the CDC has estimated that 25-40 million cases of *Streptococcus pyogenes*-induced pharyngitis occur every year, costing over $2 billion for physician visits, culture work and antibiotic therapy. There also has been an increase in toxic-shock like syndrome caused by the organism. Presently, 10,000-15,000 cases of Streptococcal and Staphylococcal Toxic shock like infections occur annually in the United States. While presently GAS infections are treated with antibiotics, given what is known about other bacteria including pneumococcus (as detailed above), the proliferation of antibiotic-resistant strains is a concern.

GAS are differentiated from other streptococci by their Group A carbohydrate, a cell wall moiety containing rhamnose and N-acetyl glucosamine. Different strains of GAS are classified, serologically, based on their M protein or on the T antigen. GAS can be assigned to 80-100 different M protein groups which form the principal basis for characterizing pathological strains. The M protein is a surface protein and is both a major virulence factor and a major protective antigen. Lancefield, *J. Immunol. 89*:307, 1962. Antibodies against M protein are opsonic and promote killing of the bacteria by phagocytes. Lancefield, *supra*.

While M proteins are potentially useful in the constitution of a vaccine, several obstacles remain on the route to an effective vaccine. First, the M protein contains epitopes that cross-react with human tissue, especially the myocardium. Dale and Beachy, *J. Exp. Med. 161*:113, 1985. Thus, anti-M protein antibodies may cause disease rather than preventing it. Second, it may not be practical to produce a vaccine against all 80-100 different strains of GAS. Any vaccine containing only a few types of M protein may be only partially effective. While the first problem might be overcome by using M protein fragments that lack the cross-reactive epitopes as immunogens (Dale et al., *J. Immunol. 151*:2188-94, 1993), such an approach has not yet been proven, and the latter problem of immunizing against numerous distinct M proteins still needs to be overcome. Accordingly, there is a need in the art for a composition which provides

generates an immunogenic response to *S. pyogenes* that is not based on the antigenicity of the M proteins.

SUMMARY OF THE INVENTION

The present invention provides methods and compositions comprising isolated nucleic acid molecules specific to *Streptococcus pneumoniae* and *Streptococcus pyogenes*, as well as vector constructs and isolated polypeptides specific to *Streptococcus pneumoniae* and *Streptococcus pyogenes*. Such compositions and methods are useful for the diagnosis of Streptococcal infection and for generating an immune response to Streptococcal bacteria.

Thus, in one aspect the present invention provides an isolated nucleic acid molecule encoding a *Streptococcus pneumoniae* Hsp60 and/or a *Streptococcus pyogenes* Hsp60. In some embodiments, the isolated nucleotide molecule is selected from the group consisting of: (a) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:1 from nucleotides 15-1652; (b) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:3 from nucleotides 15-1640; (c) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:5 from nucleotides 15-1649; (d) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:7 from nucleotides 15-1652; (e) an isolated nucleic acid molecule complementary to any one of the nucleotides of SEQ ID NOS:1, 3, 5 or 7 set forth in (a) through (d), respectively; (f) an isolated nucleic acid molecule that hybridizes under conditions of high stringency to the nucleic acid molecules of any one of (a) through (e).

In another aspect in one aspect the present invention provides an isolated nucleic acid molecule that specifically hybridizes to the nucleic acid molecule of any one of SEQ ID NO:1 from nucleotides 15-1652, SEQ ID NO:3 from nucleotides 15-1640, SEQ ID NO:5 from nucleotides 15-1649, or SEQ ID NO:7 from nucleotides 15-1652 or a complement thereof under conditions of high stringency. In further aspects the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that is identical to a segment comprising at least 25% of contiguous

nucleotide bases of any one of SEQ ID NO:1 from nucleotides 15-1652, SEQ ID NO:3 from nucleotides 15-1640, SEQ ID NO:5 from nucleotides 15-1649, or SEQ ID NO:7 from nucleotides 15-1652 or a complement thereof or an isolated nucleic acid molecule encoding Hsp60 comprising a nucleic acid sequence that encodes a polypeptide comprising any one of SEQ ID NOS: 2, 4, 6 or 8 or a variant Hsp60 that is at least 95% homologous to a polypeptide according to any one of SEQ ID NOS: 2, 4, 6 or 8.

In one embodiment, the present invention provides an isolated nucleic acid molecule according as described above, the molecule encoding a polypeptide that is able to be selectively bound by an antibody specific for a *Streptococcus pneumoniae* Hsp60 or a *Streptococcus pyogenes* Hsp60.

In still another aspect in one aspect the present invention provides an isolated nucleic acid molecule encoding at least 8 amino acids of a Streptococcal Hsp60 polypeptide selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-541 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, wherein the encoded Streptococcal Hsp60 polypeptide is able to bind to a major histocompatibility complex.

In still further aspects the present invention provides an isolated *Streptococcus pneumoniae* Hsp60 polypeptide and an isolated *Streptococcus pyogenes* Hsp60 polypeptide.

In some embodiments, the isolated Hsp60 polypeptide comprises the amino acid sequence of any one of a Streptococcal Hsp60 polypeptide selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-541 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, or variants thereof, preferably wherein the polypeptide is able to be selectively bound by an antibody specific for either a *Streptococcus pneumoniae* Hsp60 and/or *Streptococcus pyogenes* Hsp60. In further embodiments, the isolated Hsp60 polypeptide is fused to an additional polypeptide to create a fusion protein.

In still yet further aspects the present invention provides an isolated Hsp60 polypeptide comprising at least 8 amino acids selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-541 of SEQ ID NO:4, amino acid

residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, wherein the Hsp60 polypeptide is capable of binding to a major histocompatibility complex and eliciting or enhancing an immune response to *Streptococcus* in a human being.

In certain embodiments, the isolated Hsp60 polypeptide is derived from proteolytic cleavage or chemical synthesis, is an expression product of a transformed host cell containing a nucleic acid molecule encoding the Hsp60 or portion thereof. In further certain embodiments, the isolated Hsp60 polypeptide comprises greater than 95% homology to any one of a Streptococcal Hsp60 polypeptide selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-5410 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, and wherein the Hsp60 polypeptide is able to be selectively bound by an antibody specific for either a *Streptococcus pneumoniae* Hsp60 or *Streptococcus pyogenes* Hsp60 or both.

In still yet another aspect the present invention provides an isolated polypeptide wherein the polypeptide is an expression product of a transformed host cell containing one or more of the nucleic acid molecules described herein.

In still yet further aspects the present invention provides vectors comprising one or more of the nucleic acid molecules described herein. In certain embodiments, the vector is an expression vector comprising a promoter in operative linkage with the isolated nucleic acid molecule encoding the Hsp60 or portion thereof, preferably further comprising a selectable or identifiable marker and/or wherein the promoter is a constitutive or an inducible promoter. The present invention also provides host cells containing such vectors. In certain embodiments, the host cell is selected from the group consisting of a bacterial cell, a mammalian cell, a yeast cell and an insect cell.

In still yet other aspects the present invention provides compositions comprising an Hsp60 polypeptide as described herein in combination with a pharmaceutically acceptable carrier or diluent. In certain embodiments, the

composition is suitable for systemic administration, oral administration or parenteral administration.

In yet other aspects the present invention provides methods for eliciting or enhancing an immune response in a mammal against *Streptococcus*, comprising administering to the mammal an effective amount of an Hsp60 polypeptide as described herein in combination with a pharmaceutically acceptable carrier or diluent, methods for eliciting or enhancing an immune response in a mammal against a target antigen comprising administering to the mammal the target antigen joined to an Hsp60 polypeptide as described herein in combination with a pharmaceutically acceptable carrier or diluent.

In another aspect the present invention provides compositions comprising an isolated nucleic acid molecule as described herein wherein the isolated nucleic acid molecule encodes a polypeptide having at least one amino acid difference from a corresponding polypeptide of an Hsp60 protein from an organism other than *Streptococcus*.

These and other aspects of the present invention will become evident upon reference to the present specification and the attached drawings. In addition, various references are set forth herein that describe in more detail certain procedures or compositions (*e.g.*, plasmids, etc.); all such references are incorporated herein by reference in their entirety.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts the nucleotide and amino acid sequences of *Streptococcus pneumoniae* Hsp60-1 gene (SEQ ID NOS: 1 and 2 respectively).

Figure 2 depicts the nucleotide and amino acid sequences of *Streptococcus pneumoniae* Hsp60-2 gene (SEQ ID NOS: 3 and 4 respectively).

Figure 3 depicts the nucleotide and amino acid sequences of *Streptococcus pyogenes* Hsp60-1 gene (SEQ ID NOS: 5 and 6 respectively).

Figure 4 depicts the nucleotide and amino acid sequences of *Streptococcus pyogenes* Hsp60-2 gene (SEQ ID NOS: 7 and 8 respectively).

Figure 5 is a schematic representation of the sequencing strategy used to deduce the sequences of the Hsp60 genes from *S. pneumoniae* and *S. pyogenes.*

Figures 6-9 depict maps of expression vectors pETP60-1, pETP60-2, pETY60-1, and pETY60-2, which vectors include the Hsp60 genes from *S. pneumoniae* or *S. pyogenes*, respectively.

Figure 10 depicts a comparison of the *S. pneumoniae* (SEQ ID NOS: 2 and 4) and *S. pyogenes* (SEQ ID NOS: 6 and 8) Hsp60 genes with similar genes from other organisms (SEQ ID NOS: 9 through 34).

Figure 11 depicts RP-HPLC chromatograms of the Hsp60 genes from *S. pneumoniae* and *S. pyogenes.*

DETAILED DESCRIPTION OF INVENTION

The present invention provides methods and compositions comprising isolated nucleic acid molecules and polypeptides specific to *Streptococcus pneumoniae* and *Streptococcus pyogenes*, as well as vector constructs, antibodies and other materials related to isolated nucleic acid molecules and polypeptides. Such compositions and methods are useful for the diagnosis of Streptococcal infection and for generating an immune response to Streptococcal bacteria.

A "stress gene," also known as "heat shock gene," is a gene that is activated or otherwise detectably upregulated due to the contact or exposure of an organism (containing the gene) to a stressor, such as heat shock or glucose deprivation or glucose addition. A given "stress gene" also includes homologous genes within known stress gene families, such as certain genes within the Hsp60, Hsp70 and Hsp90 stress gene families, even though such homologous genes are not themselves induced by a stressor. As defined herein, a "stress protein," also known as a "heat shock protein," ("Hsp") is a protein that is encoded by a stress gene, and is therefore typically produced in significantly greater amounts upon the contact or exposure to the stressor of the organism. Each of the terms stress gene and stress protein as used in the present specification are inclusive of the other, unless the context indicates otherwise. Streptococcal Hsps, as well as Hsps from other organisms, appear to participate in

important cellular processes such as protein synthesis and assembly and disassembly of protein complexes.

A variety of stress genes and proteins are well known in the art and include, for example, Hsp100-200, Hsp100, Hsp90, Lon, Hsp70, Hsp60, TF55, Hsp40, FKBPs, cyclophilins, Hsp20-30, ClpP, GrpE, Hsp10, ubiquitin, calnexin, peptidyl-prolyl cis-trans isomerases, and protein disulfide isomerases. Macario, A.J.L., *Int. J. Clin. Lab. Res. 25*:59-70, 1995; Parsell, D.A., & Lindquist, S., *Ann. Rev. Genet. 27*: 437-496 (1993); U.S. Patent No. 5,232,833 (Sanders et al.).

In bacteria, the predominant stress proteins are proteins with molecular sizes of about 60 and 70 kDa (*i.e.*, Hsp60 and Hsp70, respectively). Hsp70 and Hsp60 typically represent about 1-3% of bacterial cell protein based on the staining pattern using sodium dodecyl sulfate-polyacrylamide gel electrophoresis ("SDS-PAGE") and the stain coomassie blue, but accumulate to levels as high as 25% under stressful conditions. Thus, Hsps are produced in an invading bacterium due to stresses put on the bacterium by the environment of the animal, and the Hsps become some of the most significant bacterial antigens displayed to the host and to which the host mounts an immune response. Therefore, by administering a Streptococcal Hsp to an animal, the Streptococcal Hsp can induce an immune response in the animal to *Streptococcus*, preferably providing resistance to such a bacterial infection. Accordingly, the isolation of Streptococcal Hsp60 genes provides a platform for the generation of isolated polypeptides or fragments or variants of Streptococcal Hsp60 useful in diagnosis and inhibition of Streptococcal associated disorders.

As used herein, "polypeptide" refers to full length proteins and fragments thereof.

As used herein, "peptide" refers to a fragment of the whole protein, whether chemically or biologically produced.

As used herein, "immunogenic" refers to an antigen or composition that elicits an immune response.

An "isolated nucleic acid molecule" refers to a polynucleotide molecule, in the form of a separate fragment or as a component of a larger nucleic acid construct,

that has been separated from its source cell (including the chromosome it normally resides in) at least once in a substantially pure form. Nucleic acid molecules can be comprised of a wide variety of nucleotides and molecules well known in the art, including DNA, RNA, nucleic acid analogues, or any combination of these.

As used herein, "vector" refers to a polynucleotide assembly capable of directing expression and/or replication of the nucleic acid sequence of interest. Such assembly can, if desired, be included as a part of other components, such as a protein, lipid or lipoprotein coat, for delivery of the vector or for other purposes.

An "expression vector" refers to polynucleotide vector having at least a promoter sequence operably linked to the nucleic acid sequence of interest.

As used herein, a "promoter" refers to a nucleotide sequence that contains elements that direct the transcription of an operably linked nucleic acid sequence. At minimum, a promoter contains an RNA polymerase binding site. Promoter regions can also contain enhancer elements which by definition enhance transcription.

A. HSP60 GENES AND POLYPEPTIDES FROM *STREPTOCOCCUS PNEUMONIAE* AND *STREPTOCOCCUS PYOGENES*

As used herein, Hsp60 refers to heat shock genes from the Hsp60 family of genes that produce heat shock proteins of approximately 60kDa; the nucleotide and amino acid sequences of Hsp60 genes and gene products from *Streptococcus pneumoniae* and *Streptococcus pyogenes* are set forth in Figures 1-4 (SEQ ID NOS:1-8; such sequences also include the PCR primers used to isolate the Hsp60 genes). Within the context of this invention it should be understood that Hsp60 includes wild-type/native protein sequences, as well as other variants (including alleles) and fragments of the native protein sequence. Briefly, such variants may result from natural polymorphisms or be synthesized by recombinant methodology or chemical synthesis, and differ from wild-type proteins by one or more amino acid substitutions, insertions, deletions, or the like. Further, in the region of homology to the native sequence, variants should preferably have at least 95% amino acid sequence homology, and within

certain embodiments, greater than 97% or 98% homology. As will be appreciated by those of ordinary skill in the art, a nucleotide sequence encoding Hsp60 or variant may differ from the native sequences presented herein due to codon degeneracies, nucleotide polymorphisms, or nucleotide substitutions, deletions or insertions.

An "isolated nucleic acid molecule encoding *Streptococcus* Hsp60" refers to nucleic acid sequences that are capable of encoding Hsp60 polypeptides of *Streptococcus*, preferably *Streptococcus pneumoniae* or *Streptococcus pyogenes*. While several embodiments of such molecules are depicted in SEQ ID NOS:1-4, it should be understood that within the context of the present invention, reference to one or more of these genes includes variants of the genes, that is, naturally occurring variants or sequences that are substantially similar to the genes (and, where appropriate, the protein (including peptides and polypeptides) that are encoded by the genes and their variants). As used herein, the nucleotide sequence is deemed to be "substantially similar" if: (a) the nucleotide sequence is derived from the coding region of a native gene of *Streptococcus* and maintains substantially the same biological activity (including, for example, portions of the sequence or allelic variations of the sequences discussed above); or (b) the nucleotide sequence is capable of hybridization to the nucleotide sequences of the present invention under high stringency (*i.e.*, capable of selectively hybridizing to nucleotide sequences from *Streptococcus*); or (c) the nucleotide sequences are degenerate (*i.e.*, sequences which code for the same amino acid using a different codon sequence) as a result of the genetic code to the nucleotide sequences defined in (a) or (b); or (d) is a complement of any of the sequences described in (a), (b) or (c)

One aspect of the present invention is the use of *Streptococcus* Hsp60 nucleotide sequences to produce recombinant proteins for immunizing an animal. Therefore, the use of any length of nucleic acid disclosed by the present invention (preferably 24 nucleotides or longer) which encodes a polypeptide or fragment thereof that is capable of binding to the major histocompatibility complex and eliciting or enhancing an immunogenic response is contemplated by this invention. Immunogenic response can be readily tested by known methods such as challenging a mouse or rabbit

with the antigen of interest and thereafter collecting plasma and determining if the antibody of interest is present. Other assays particularly useful for the detection of T-cell responses include proliferation assays, T-cell cytotoxicity assays and assays for delayed hypersensitivity. In determining whether an antibody specific for the antigen of interest was produced by the animal, many diagnostic tools are available, for example, testing binding of labeled antigen to plasma derived antibodies, or using Enzyme-linked immunoassays with tag attached to the antigen of interest.

The Streptococcal Hsp60 genes of this invention can be obtained using a variety of methods. For example, a nucleic acid molecule can be obtained from a cDNA or genomic expression library by screening with an antibody or antibodies reactive to one or more of these Hsp60s (*see, e.g.*, Sambrook et al., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor, 1989; Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing, 1987). Further, random-primed PCR can be employed (*see, e.g., Methods in Enzymol. 254*:275, 1995). In addition, variations of random-primed PCR can also be used, especially when a particular gene or gene family is desired. In one such method, one of the primers is a poly deoxy-thymine and the other is a degenerate primer based on the amino acid sequence or nucleotide sequence of related Hsps.

Other methods can also be used to obtain a nucleic acid molecule that encodes Streptococcal Hsp60. For example, a nucleic acid molecule can be obtained by using the sequence information provided herein to synthesize a probe which can be labeled, such as with a radioactive label, enzymatic label, protein label, fluorescent label, or the like, and hybridized to a genomic library or a cDNA library constructed in a phage, plasmid, phagemid, or other viral vector (*see, e.g.*, Sambrook et al. (*supra*); Ausubel et al. (*supra*)). DNA representing RNA or genomic nucleic acid sequence can also be obtained by amplification using sets of primers complementary to 5' and 3' sequences of the cDNA sequence, such as presented in Example 1. For ease of cloning, restriction sites can also be incorporated into the primers.

Variants (including alleles) of the Hsp60 genes provided herein can be readily isolated from natural variants (*e.g.*, polymorphisms, mutants), synthesized or

constructed. Many methods have been developed for generating mutants (*see generally* Sambrook et al. (*supra*); Ausubel et al. (*supra*)). Briefly, preferred methods for generating nucleotide substitutions utilize an oligonucleotide that spans the base or bases to be mutated and contains the mutated base or bases. The oligonucleotide is hybridized to complementary single stranded nucleic acid and second strand synthesis is primed from the oligonucleotide. The double-stranded nucleic acid is prepared for transformation into host cells, such as *E. coli*, other prokaryotes, yeast or other eukaryotes. Standard screening and vector growth protocols are used to identify mutant sequences and obtain high yields.

Similarly, deletions and/or insertions of the Hsp60 gene can be constructed by any of a variety of known methods. For example, the gene can be digested with restriction enzymes and religated such that sequence is deleted or religated with additional sequence such that an insertion or large substitution is made. Other means of generating variant sequences, known in the art, can be employed, for examples see Sambrook et al. (*supra*) and Ausubel et al. (*supra*). Moreover, verification of variant sequences is typically accomplished by restriction enzyme mapping, sequence analysis or hybridization. Variants which encode a polypeptide that elicits an immunogenic response specific for *Streptococcus* are useful in the context of this invention.

As noted above, the present invention also provides isolated polypeptides. Within the context of the present invention, unless otherwise clear from the context, such polypeptides are understood to include the whole, or portions/fragments, of a gene product derived from one or more of the Streptococcal Hsp60 genes or derivatives thereof as discussed above. In one aspect of the present invention, the protein is encoded by a portion of a native gene or is encoded by a derivative of a native gene and the protein or fragment thereof elicits or enhances an immune response specific for *Streptococcus*.

A "purified" Hsp60 stress protein of the present invention is a heat shock protein of the Hsp60 family from *Streptococcus pneumoniae* or *Streptococcus pyogenes* that has been purified from its producing cell. For example, the Streptococcal Hsp60

polypeptides of the present invention can be purified by a variety of standard methods with or without a detergent purification step. For example, Streptococcal Hsp60 can be isolated by, among other methods, culturing suitable host and vector systems to produce recombinant Hsp60 (discussed further herein). Then, supernatants from such cell lines, or Hsp60 inclusions, or whole cells where the Hsp60 is not excreted into the supernatant, can be treated by a variety of purification procedures. For example, the Streptococcal Hsp60-containing composition can be applied to a suitable purification matrix such as an anti-Hsp60 antibody bound to a suitable support. Alternatively, anion or cation exchange resins, gel filtration or affinity, hydrophobic or reverse phase chromatography may be employed in order to purify the protein. The Hsp60 polypeptide can also be concentrated using commercially available protein concentration filters, such as an Amicon or Millipore Pellicon ultrafiltration unit, or by vacuum dialysis. In another alternative the supernatant can first be concentrated using one of the above mentioned protein concentration filters, followed by application of the concentrate to a suitable purification matrix such as those described above.

In one embodiment, the isolated Streptococcal Hsp60s of the present invention are produced in a recombinant form, utilizing genetic manipulation techniques that are well known in the art. For example, Streptococcal Hsp60 can be expressed as a histidine-tagged molecule, permitting purification on a nickel-chelating matrix. Alternatively, other tags may be used, including FLAG and GST. The associated tag can then be removed in the last step of purification, for example, for certain vectors, His-tagged proteins may be incubated with thrombin, resulting in cleavage of a recognition sequence between the tag and the Hsp60 polypeptide (*e.g.*, pET vectors from Invitrogen). Following purification of Streptococcal Hsp60 from a gram-negative bacterial host, whether tagged or not, it will be necessary to reduce the level of endotoxin in the Hsp60 preparation, as discussed above.

B. VECTORS, HOST CELLS, AND EXPRESSION OF STREPTOCOCCAL HSP60

It is well known in the art that certain vectors (*e.g.*, pUC) can be used for producing multiple copies of a nucleotide molecule of interest as well as being useful

for genetic manipulation techniques (*e.g.*, site-directed mutagenesis). See Sambrook (*supra*). Of particular interest to this disclosure are expression vectors. The expression vector includes transcriptional promoter/enhancer elements operably linked to the Streptococcal Hsp60 nucleic acid molecule. The expression vector may be composed of either deoxyribonucleic acids ("DNA"), ribonucleic acids ("RNA"), or a combination of the two (*e.g.*, a DNA-RNA chimera). Optionally, the expression vector may include a polyadenylation sequence or one or more restriction sites. Additionally, depending on the host cell chosen and the expression vector employed, other genetic elements such as an origin of replication, additional nucleic acid restriction sites, enhancers, sequences conferring inducibility of transcription, and genes encoding proteins suitable for use as selectable or identifiable markers, may also be incorporated into the expression vectors described herein.

The manipulation and expression of Streptococcal Hsp60 genes can be accomplished by culturing host cells containing an expression vector capable of expressing the Hsp60 genes. Such vectors or vector constructs include either synthetic or cDNA-derived nucleic acid molecules or genomic DNA fragments encoding Streptococcal Hsp60 polypeptides, which are operably linked to suitable transcriptional or translational regulatory elements. Suitable regulatory elements within the expression vector can be derived from a variety of sources, including bacterial, fungal, viral, mammalian, insect, or plant genes. Selection of appropriate regulatory elements is dependent on the host cell chosen, and can be readily accomplished by one of ordinary skill in the art in light of the present specification. Examples of regulatory elements include a transcriptional promoter and enhancer or RNA polymerase binding sequence, a transcriptional terminator, and a ribosomal binding sequence, including a translation initiation signal.

Nucleic acid molecules that encode any of the Streptococcal Hsp60 polypeptides described above can be expressed by a wide variety of prokaryotic and eukaryotic host cells, including bacterial, mammalian, yeast or other fungi, viral, insect, and plant cells. The selection of a host cell may also assist the production of glycosolated or non-glycosolated Hsp60s, depending upon the desires of the user.

Methods for transforming or transfecting such cells to express nucleic acids are well known in the art (*see, e.g.*, Itakura et al., U.S. Patent No. 4,704,362; Hinnen et al., *PNAS* USA *75*:1929-1933, 1978; Murray et al., U.S. Patent No. 4,801,542; Upshall et al., U.S. Patent No. 4,935,349; Hagen et al., U.S. Patent No. 4,784,950; Axel et al., U.S. Patent No. 4,399,216; Goeddel et al., U.S. Patent No. 4,766,075; and Sambrook et al., *Molecular Cloning: A Laboratory Manual, 2nd edition*, Cold Spring Harbor Laboratory Press, 1989; for plant cells see Czako and Marton, *Plant Physiol. 104*:1067-1071, 1994; Paszkowski et al., *Biotech. 24*:387-392, 1992).

Bacterial host cells suitable for carrying out the present invention include *E. coli*, such as *E. coli* DH5α (Stratagene, La Jolla, California), *M. leprae, M. tuberculosis, M. bovis, B. subtilis, Salmonella typhimurium*, and various species within the genera *Pseudomonas, Streptomyces, Streptococcus*, and *Staphylococcus*, as well as many other bacterial species well known to one of ordinary skill in the art.

Bacterial expression vectors preferably comprise a promoter, which functions in the host cell, one or more selectable phenotypic markers, and a bacterial origin of replication. Representative promoters include the β-lactamase (penicillinase) and lactose promoter system (see Chang et al., *Nature 275*:615, 1978), the T7 RNA polymerase promoter (Studier et al., *Meth. Enzymol. 185*:60-89, 1990), the lambda promoter (Elvin et al., *Gene 87*:123-126, 1990), the *trp* promoter (Nichols and Yanofsky, *Meth. in Enzymology 101*:155, 1983) and the *tac* promoter (Russell et al., *Gene 20*: 231, 1982). Representative selectable markers include various antibiotic resistance markers such as the kanamycin or ampicillin resistance genes. Many plasmids suitable for transforming host cells are well known in the art, including among others, pBR322 (*see* Bolivar et al., *Gene 2*:95, 1977), the pUC plasmids pUC18, pUC19, pUC118, pUC119 (*see* Messing, *Meth. in Enzymology 101*:20-77, 1983; Vieira and Messing, *Gene 19*:259-268, 1982), and pNH8A, pNH16a, pNH18a, and Bluescript M13 (Stratagene, La Jolla, Calif.).

Fungal host cells suitable for carrying out the present invention include, among others, *Saccharomyces pombe, Saccharomyces cerevisiae*, the genera *Pichia* or *Kluyveromyces* and various species of the genus *Aspergillus* (McKnight et al., U.S.

Patent No. 4,935,349). Suitable expression vectors for yeast and fungi include, among others, YCp50 (ATCC No. 37419) for yeast, and the amdS cloning vector pV3 (Turnbull, *Bio/Technology 7*:169, 1989), YRp7 (Struhl et al., *Proc. Natl. Acad. Sci. USA 76*:1035-1039, 1978), YEp13 (Broach et al., *Gene 8*:121-133, 1979), pJDB249 and pJDB219 (Beggs, *Nature 275*:104-108, 1978) and derivatives thereof.

Preferred promoters for use in yeast include promoters from yeast glycolytic genes (Hitzeman et al., *J. Biol. Chem. 255*:12073-12080, 1980; Alber and Kawasaki, *J. Mol. Appl. Genet. 1*:419-434, 1982) or alcohol dehydrogenase genes (Young et al., in *Genetic Engineering of Microorganisms for Chemicals*, Hollaender et al. (eds.), p. 355, Plenum, New York, 1982; Ammerer, *Meth. Enzymol. 101*:192-201, 1983). Examples of useful promoters for fungi vectors include those derived from *Aspergillus nidulans* glycolytic genes, such as the *adh3* promoter (McKnight et al., *EMBO J. 4*:2093-2099, 1985). The expression units may also include a transcriptional terminator. An example of a suitable terminator is the *adh3* terminator (McKnight et al., *ibid.*, 1985).

As with bacterial vectors, the yeast vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers include those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include leu2 (Broach et al., *ibid.*), ura3 (Botstein et al., *Gene 8*:17, 1979), or *his3* (Struhl et al., *ibid.*). Another suitable selectable marker is the *cat* gene, which confers chloramphenicol resistance on yeast cells.

Techniques for transforming fungi are well known in the literature, and have been described, for instance, by Beggs (*ibid.*), Hinnen et al.(*Proc. Natl. Acad. Sci. USA 75*:1929-1933, 1978), Yelton et al. (*Proc. Natl. Acad. Sci. USA 81*:1740-1747, 1984), and Russell (*Nature 301*:167-169, 1983). The genotype of the host cell may contain a genetic defect that is complemented by the selectable marker present on the expression vector. Choice of a particular host and selectable marker is well within the level of ordinary skill in the art in light of the present specification.

Protocols for the transformation of yeast are also well known to those of ordinary skill in the art. For example, transformation may be readily accomplished either by preparation of spheroplasts of yeast with DNA (*see* Hinnen et al., *PNAS USA 75*:1929, 1978) or by treatment with alkaline salts such as LiCl (*see* Itoh et al., *J. Bacteriology 153*:163, 1983). Transformation of fungi may also be carried out using polyethylene glycol as described by Cullen et al. (*Bio/Technology 5*:369, 1987).

Viral vectors include those that comprise a promoter that directs the expression of an isolated nucleic acid molecule that encodes a Streptococcal Hsp60 as described above. A wide variety of promoters may be utilized within the context of the present invention, including for example, promoters such as MoMLV LTR, RSV LTR, Friend MuLV LTR, adenoviral promoter (Ohno et al., *Science 265*: 781-784, 1994), neomycin phosphotransferase promoter/enhancer, late parvovirus promoter (Koering et al., *Hum. Gene Therap. 5*:457-463, 1994), Herpes TK promoter, SV40 promoter, metallothionein IIa gene enhancer/promoter, cytomegalovirus immediate early promoter, and the cytomegalovirus immediate late promoter. The promoter may also be a tissue-specific promoter (*see e.g.,* WO 91/02805; EP 0,415,731; and WO 90/07936). In addition to the above-noted promoters, other viral-specific promoters (*e.g.*, retroviral promoters (including those noted above, as well as others such as HIV promoters), hepatitis, herpes (*e.g.,* EBV), and bacterial, fungal or parasitic-specific (*e.g.,* malarial-specific) promoters may be utilized in order to target a specific cell or tissue which is infected with a virus, bacteria, fungus or parasite.

Thus, Streptococcal Hsp60 polypeptides of the present invention may be expressed from a variety of viral vectors, including for example, herpes viral vectors (*e.g.*, U.S. Patent No. 5,288,641), adenoviral vectors (*e.g.*, WO 94/26914, WO 93/9191; Kolls et al., *PNAS 91*(1):215-219, 1994; Kass-Eisler et al., *PNAS 90*(24):11498-502, 1993; Guzman et al., *Circulation 88*(6):2838-48, 1993; Guzman et al., *Cir. Res. 73*(6):1202-1207, 1993; Zabner et al., *Cell 75*(2):207-216, 1993; Li et al., *Hum Gene Ther. 4*(4):403-409, 1993; Caillaud et al., *Eur. J. Neurosci. 5*(10):1287-1291, 1993; Vincent et al., *Nat. Genet. 5*(2):130-134, 1993; Jaffe et al., *Nat. Genet. 1*(5):372-378, 1992; and Levrero et al., *Gene 101*(2):195-202, 1991), adenovirus-associated viral

vectors (Flotte et al., *PNAS 90*(22):10613-10617, 1993), baculovirus vectors, parvovirus vectors (Koering et al., *Hum. Gene Therap. 5*:457-463, 1994), pox virus vectors (Panicali and Paoletti, *PNAS 79*:4927-4931, 1982; and Ozaki et al., *Biochem. Biophys. Res. Comm. 193*(2):653-660, 1993), and retroviruses (*e.g.*, EP 0,415,731; WO 90/07936; WO 91/0285, WO 94/03622; WO 93/25698; WO 93/25234; U.S. Patent No. 5,219,740; WO 93/11230; WO 93/10218. Within various embodiments, either the viral vector itself or a viral particle which contains the viral vector may be utilized in the methods and compositions described below.

Mammalian cells suitable for carrying out the present invention include, among others: PC12 (ATCC No. CRL1721), N1E-115 neuroblastoma, SK-N-BE(2)C neuroblastoma, SHSY5 adrenergic neuroblastoma, NS20Y and NG108-15 murine cholinergic cell lines, or rat F2 dorsal root ganglion line, COS (*e.g.*, ATCC No. CRL 1650 or 1651), BHK (*e.g.*, ATCC No. CRL 6281; BHK 570 cell line (deposited with the American Type Culture Collection under accession number CRL 10314), CHO (ATCC No. CCL 61), HeLa (*e.g.*, ATCC No. CCL 2), 293 (ATCC No. 1573; Graham et al., *J. Gen. Virol. 36*:59-72, 1977) and NS-1 cells. Other mammalian cell lines may be used within the present invention, including Rat Hep I (ATCC No. CRL 1600), Rat Hep II (ATCC No. CRL 1548), TCMK (ATCC No. CCL 139), Human lung (ATCC No. CCL 75.1), Human hepatoma (ATCC No. HTB-52), Hep G2 (ATCC No. HB 8065), Mouse liver (ATCC No. CCL 29.1), NCTC 1469 (ATCC No. CCL 9.1), SP2/0-Ag14 (ATCC No. 1581), HIT-T15 (ATCC No. CRL 1777), and RINm 5AHT2B (Orskov and Nielson, *FEBS 229*(1):175-178, 1988).

Mammalian expression vectors for use in carrying out the present invention include a promoter capable of directing the transcription of a cloned gene or cDNA. Preferred promoters include viral promoters and cellular promoters. Viral promoters include the cytomegalovirus immediate early promoter (Boshart et al., *Cell 41*:521-530, 1985), cytomegalovirus immediate late promoter, SV40 promoter (Subramani et al., *Mol. Cell. Biol. 1*:854-864, 1981), MMTV LTR, RSV LTR, metallothionein-1, adenovirus E1a. Cellular promoters include the mouse metallothionein-1 promoter (Palmiter et al., U.S. Patent No. 4,579,821), action

promoters, a mouse V_H promoter (Bergman et al., *Proc. Natl. Acad. Sci. USA 81*:7041-7045, 1983; Grant et al., *Nucl. Acids Res. 15*:5496, 1987) and a mouse V_H promoter (Loh et al., *Cell 33*:85-93, 1983). The choice of promoter will depend, at least in part, upon the level of expression desired or the recipient cell line to be transfected.

Such expression vectors can also contain a set of RNA splice sites located downstream from the promoter and upstream from the DNA sequence encoding the peptide or protein of interest. Preferred RNA splice sites may be obtained from adenovirus and/or immunoglobulin genes. Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Suitable polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman and Sharp, *ibid.*), the polyadenylation signal from the Adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., *Nuc. Acids Res. 9*:3719-3730, 1981). The expression vectors may include a noncoding viral leader sequence, such as the Adenovirus 2 tripartite leader, located between the promoter and the RNA splice sites. Preferred vectors may also include enhancer sequences, such as the SV40 enhancer. Expression vectors may also include sequences encoding the adenovirus VA RNAs. Suitable expression vectors can be obtained from commercial sources (*e.g.*, Stratagene, La Jolla, Calif.).

Vector constructs comprising cloned DNA sequences can be introduced into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., *Cell 14*:725, 1978; Corsaro and Pearson, *Somatic Cell Genetics 7*:603, 1981; Graham and Van der Eb, *Virology 52*:456, 1973), electroporation (Neumann et al., *EMBO J. 1*:841-845, 1982), or DEAE-dextran mediated transfection (Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., NY, 1987). *See generally* Sambrook et al. (supra). To identify cells that have stably integrated the cloned DNA, a selectable marker is generally introduced into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. Preferred amplifiable selectable markers are the DHFR

gene and the neomycin resistance gene. Selectable markers are reviewed by Thilly (*Mammalian Cell Technology*, Butterworth Publishers, Stoneham, MA).

Mammalian cells containing a suitable vector are allowed to grow for a period of time, typically 1-2 days, to begin expressing the DNA sequence(s) of interest. Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an amplifiable, selectable marker the drug concentration may be increased in a stepwise manner to select for increased copy number of the cloned sequences, thereby increasing expression levels. Cells expressing the introduced sequences are selected and screened for production of the protein of interest in the desired form or at the desired level. Cells that satisfy these criteria can then be cloned and scaled up for production.

Numerous insect host cells known in the art can also be useful within the present invention, in light of the subject specification. For example, the use of baculoviruses as vectors for expressing heterologous DNA sequences in insect cells has been reviewed by Atkinson et al. (*Pestic. Sci. 28*:215-224,1990).

Numerous plant host cells known in the art can also be useful within the present invention, in light of the subject specification. For example, the use of *Agrobacterium rhizogenes* as vectors for expressing genes in plant cells has been reviewed by Sinkar et al., *J. Biosci. (Bangalore) 11*:47-58, 1987.

Upon expression of the Streptococcal Hsp60 polypeptides or fragments thereof in the host cells, the polypeptide or peptide may be preliminarily released and/or isolated from the host cell utilizing methods such as those discussed previously herein.

As noted above, depending on the host cell in which one desires to express Hsp60, the gene encoding the protein is introduced into an expression vector comprising a promoter that is active in the host cell. Other components of the expression unit such as transcribed but not translated sequences at the ends of the coding region may also be selected according to the particular host utilized. In some cases, it may be necessary to introduce artificially an intervening sequence to ensure high level expression. Expression can be monitored by SDS-PAGE and staining, if expression levels are sufficiently high. Additionally, if the protein is produced with a

tag, detection by anti-tag antibody can be carried out and if produced with no tag, detection by anti-Hsp60 antibody that does not recognize homologous proteins of the host may be employed. Further, any method known in the art for protein identification may be utilized to this end (*e.g.*, a high resolution electrophoretic method or 2D electrophoresis).

C. PREPARATION OF ANTIBODIES AGAINST THE HSP60 POLYPEPTIDES OF THE PRESENT INVENTION

In another aspect, the proteins of the present invention are utilized to prepare specifically binding antibodies (*i.e.*, binding partners). Accordingly, the present invention also provides such antibodies. Within the context of the present invention, the term "antibodies" includes polyclonal antibodies, monoclonal antibodies, anti-idiotypic antibodies, fragments thereof such as $F(ab')_2$ and Fab fragments, and recombinantly or synthetically produced binding partners. Such binding partners incorporate the variable regions that permit a monoclonal antibody to specifically bind, which means an antibody able to selectively bind to a peptide produced from one of the Streptococcal Hsp60 genes of the invention. The affinity of a monoclonal antibody or binding partner can be readily determined by one of ordinary skill in the art (*see* Scatchard, *Ann. N.Y. Acad. Sci. 51*:660-672, 1949).

Polyclonal antibodies can be readily generated by one of ordinary skill in the art from a variety of warm-blooded animals such as horses, cows, goats, sheep, dogs, chickens, turkeys, rabbits, mice, or rats. Briefly, the desired protein or peptide is utilized to immunize the animal through intraperitoneal, intramuscular, intraocular, or subcutaneous injections. The immunogenicity of the protein or peptide of interest may be increased through the use of an adjuvant such as Freund's complete or incomplete adjuvant. Following several booster immunizations, small samples of serum are collected and tested for reactivity to the desired protein or peptide.

Particularly preferred polyclonal antisera give a signal that is at least three times greater than background. Once the titer of the animal has reached a plateau

in terms of its reactivity to the protein, larger quantities of polyclonal antisera may be readily obtained either by weekly bleedings, or by exsanguinating the animal.

Monoclonal antibodies can also be readily generated using well-known techniques (*see* U.S. Patent Nos. RE 32,011, 4,902,614, 4,543,439, and 4,411,993; *see also Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, Kennett, McKearn, and Bechtol (eds.), 1980, and *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988). Briefly, in one embodiment, a subject animal such as a rat or mouse is injected with a desired protein or peptide. If desired, various techniques may be utilized in order to increase the resultant immune response generated by the protein, in order to develop greater antibody reactivity. For example, the desired protein or peptide may be coupled to another protein such as ovalbumin or keyhole limpet hemocyanin (KLH), or through the use of adjuvants such as Freund's complete or incomplete adjuvant. The initial elicitation of an immune response, may preferably be through intraperitoneal, intramuscular, intraocular, or subcutaneous routes.

Between one and three weeks after the initial immunization, the animal may be reimmunized. The animal may then be test bled and the serum tested for binding to the desired antigen using assays as described above. Additional immunizations may also be accomplished until the animal has reached a plateau in its reactivity to the desired protein or peptide. The animal may then be given a final boost of the desired protein or peptide, and three to four days later sacrificed. At this time, the spleen and lymph nodes may be harvested and disrupted into a single cell suspension by passing the organs through a mesh screen or by rupturing the spleen or lymph node membranes which encapsulate the cells. Within one embodiment the red cells are subsequently lysed by the addition of a hypotonic solution, followed by immediate return to isotonicity.

Within another embodiment, suitable cells for preparing monoclonal antibodies are obtained through the use of *in vitro* immunization techniques. Briefly, an animal is sacrificed, and the spleen and lymph node cells are removed as described above. A single cell suspension is prepared, and the cells are placed into a culture

containing a form of the protein or peptide of interest that is suitable for generating an immune response as described above. Subsequently, the lymphocytes are harvested and fused as described below.

Cells that are obtained through the use of *in vitro* immunization or from an immunized animal as described above may be immortalized by transfection with a virus such as the Epstein-Barr Virus (EBV). (*See* Glasky and Reading, *Hybridoma 8*(4):377-389, 1989.) Alternatively, within a preferred embodiment, the harvested spleen and/or lymph node cell suspensions are fused with a suitable myeloma cell in order to create a "hybridoma" which secretes monoclonal antibodies. Suitable myeloma lines are preferably defective in the construction or expression of antibodies, and are additionally syngeneic with the cells from the immunized animal. Many such myeloma cell lines are well known in the art and may be obtained from sources such as the American Type Culture Collection (ATCC), Rockville, Maryland (*see Catalogue of Cell Lines & Hybridomas*, 6th ed., ATCC, 1988). Representative myeloma lines include: for humans, UC 729-6 (ATCC No. CRL 8061), MC/CAR-Z2 (ATCC No. CRL 8147), and SKO-007 (ATCC No. CRL 8033); for mice, SP2/0-Ag14 (ATCC No. CRL 1581), and P3X63Ag8 (ATCC No. TIB 9); and for rats, Y3-Ag1.2.3 (ATCC No. CRL 1631), and YB2/0 (ATCC No. CRL 1662). Particularly preferred fusion lines include NS-1 (ATCC No. TIB 18) and P3X63 - Ag 8.653 (ATCC No. CRL 1580), which may be utilized for fusions with either mouse, rat, or human cell lines. Fusion between the myeloma cell line and the cells from the immunized animal can be accomplished by a variety of methods, including the use of polyethylene glycol (PEG) (*see Antibodies: A Laboratory Manual*, Harlow and Lane, *supra*) or electrofusion. (*See* Zimmerman and Vienken, *J. Membrane Biol. 67*:165-182, 1982.)

Following the fusion, the cells are placed into culture plates containing a suitable medium, such as RPMI 1640 or DMEM (Dulbecco's Modified Eagles Medium, JRH Biosciences, Lenexa, Kan.). The medium may also contain additional ingredients, such as Fetal Bovine Serum (FBS, *e.g.*, from Hyclone, Logan, Utah, or JRH Biosciences), thymocytes that were harvested from a baby animal of the same species as was used for immunization, or agar to solidify the medium. Additionally, the medium

should contain a reagent which selectively allows for the growth of fused spleen and myeloma cells. Particularly preferred is the use of HAT medium (hypoxanthine, aminopterin, and thymidine) (Sigma Chemical Co., St. Louis, Mo.). After about seven days, the resulting fused cells or hybridomas may be screened in order to determine the presence of antibodies which recognize the desired antigen. Following several clonal dilutions and reassays, hybridoma producing antibodies that bind to the protein of interest can be isolated.

Other techniques may also be utilized to construct monoclonal antibodies. (*See* Huse et al., "Generation of a Large Combinational Library of the Immunoglobulin Repertoire in Phage Lambda," *Science 246*:1275-1281, 1989; *see also* Sastry et al., "Cloning of the Immunological Repertoire in *Escherichia coli* for Generation of Monoclonal Catalytic Antibodies: Construction of a Heavy Chain Variable Region-Specific cDNA Library," *Proc. Natl. Acad. Sci. USA 86*:5728-5732, 1989; *see also* Alting-Mees et al., "Monoclonal Antibody Expression Libraries: A Rapid Alternative to Hybridomas," *Strategies in Molecular Biology 3*:1-9, 1990; these references describe a commercial system available from Stratagene, La Jolla, California, which enables the production of antibodies through recombinant techniques.) Briefly, mRNA is isolated from a B cell population and utilized to create heavy and light chain immunoglobulin cDNA expression libraries in the λIMMUNOZAP(H) and λ IMMUNOZAP(L) vectors. These vectors may be screened individually or co-expressed to form Fab fragments or antibodies (*see* Huse et al. (*supra*); *see also* Sastry et al. (*supra*)). Positive plaques can subsequently be converted to a non-lytic plasmid which allows high level expression of monoclonal antibody fragments from *E. coli*.

Similarly, binding partners can also be constructed utilizing recombinant DNA techniques to incorporate the variable regions of a gene that encodes a specifically binding antibody. The construction of these binding partners can be readily accomplished by one of ordinary skill in the art given the disclosure provided herein. (*See* Larrick et al., "Polymerase Chain Reaction Using Mixed Primers: Cloning of Human Monoclonal Antibody Variable Region Genes From Single Hybridoma Cells,"

Biotechnology 7:934-938, 1989; Riechmann et al., "Reshaping Human Antibodies for Therapy," *Nature* 332:323-327, 1988; Roberts et al., "Generation of an Antibody with Enhanced Affinity and Specificity for its Antigen by Protein Engineering," *Nature* 328:731-734, 1987; Verhoeyen et al., "Reshaping Human Antibodies: Grafting an Antilysozyme Activity," *Science* 239:1534-1536, 1988; Chaudhary et al., "A Recombinant Immunotoxin Consisting of Two Antibody Variable Domains Fused to *Pseudomonas* Exotoxin," *Nature* 339:394-397, 1989; *see also* U.S. Patent No. 5,132,405 entitled "Biosynthetic Antibody Binding Sites.") Briefly, in one embodiment, DNA segments encoding the desired protein or peptide of interest-specific antigen binding domains are amplified from hybridomas that produce a specifically binding monoclonal antibody, and are inserted directly into the genome of a cell that produces human antibodies. (*See* Verhoeyen et al. (*supra*); *see also* Reichmann et al. (*supra*)). This technique allows the antigen-binding site of a specifically binding mouse or rat monoclonal antibody to be transferred into a human antibody. Such antibodies are preferable for therapeutic use in humans because they are not as antigenic as rat or mouse antibodies.

In an alternative embodiment, genes that encode the variable region from a hybridoma producing a monoclonal antibody of interest are amplified using oligonucleotide primers for the variable region. These primers may be synthesized by one of ordinary skill in the art, or may be purchased from commercially available sources. For instance, primers for mouse and human variable regions including, among others, primers for V_{Ha}, V_{Hb}, V_{Hc}, V_{Hd}, C_{H1}, V_L and C_L regions, are available from Stratagene (La Jolla, Calif.). These primers may be utilized to amplify heavy or light chain variable regions, which may then be inserted into vectors such as IMMUNOZAP™(H) or IMMUNOZAP™(L) (Stratagene), respectively. These vectors may then be introduced into *E. coli* for expression. Utilizing these techniques, large amounts of a single-chain polypeptide containing a fusion of the V_H and V_L domains may be produced (*see* Bird et al., *Science* 242:423-426, 1988).

Monoclonal antibodies and other binding partners can be produced in a number of host systems, including tissue cultures, bacteria, eukaryotic cells, plants and other host systems known in the art.

Once suitable antibodies or binding partners have been obtained, they may be isolated or purified by many techniques well known to those of ordinary skill in the art (*see Antibodies: A Laboratory Manual*, Harlow and Lane (*supra*)). Suitable techniques include peptide or protein affinity columns, HPLC or RP-HPLC, purification on protein A or protein G columns, or any combination of these techniques. Within the context of the present invention, the term "isolated" as used to define antibodies or binding partners means "substantially free of other blood components."

The binding partners of the present invention have many uses. For example, antibodies can be utilized in flow cytometry to identify cells bearing such a protein. Briefly, in order to detect the protein or peptide of interest on cells, the cells are incubated with a labeled monoclonal antibody which specifically binds to the protein of interest, followed by detection of the presence of bound antibody. Labels suitable for use within the present invention are well known in the art including, among others, flourescein isothiocyanate (FITC), phycoerythrin (PE), horse radish peroxidase (HRP), and colloidal gold. Particularly preferred for use in flow cytometry is FITC, which may be conjugated to purified antibody according to the method of Keltkamp in "Conjugation of Fluorescein Isothiocyanate to Antibodies. I. Experiments on the Conditions of Conjugation," *Immunology 18*:865-873, 1970. (*See also* Keltkamp, "Conjugation of Fluorescein Isothiocyanate to Antibodies. II. A Reproducible Method," *Immunology 18*:875-881, 1970; Goding, "Conjugation of Antibodies with Fluorochromes: Modification to the Standard Methods," *J. Immunol. Methods 13*:215-226, 1970.) The antibodies can also be used to target drugs to *Streptococcus* as well as a diagnostic for determining Streptococcal infection.

D. Assays That Utilize The Hsp60 Polypeptides, Or Antibodies Thereto, Of The Present Invention

A variety of assays can be utilized in order to detect the Hsp60 polypeptides from *Streptococcus pneumoniae* and *Streptococcus pyogenes* of the present invention, or antibodies that specifically bind to such Hsp60 polypeptides. Exemplary assays are described in detail in *Antibodies: A Laboratory Manual,* Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988. Representative examples of such assays include: countercurrent immuno-electrophoresis (CIEP), radioimmunoassays, radioimmunoprecipitations, enzyme-linked immuno-sorbent assays (ELISA), dot blot assays, inhibition or competition assays, and sandwich assays, immunostick (dipstick) assays, simultaneous immunoassays, immunochromatographic assays, immunofiltration assays, latex bead agglutination assays, immunofluorescent assays, biosensor assays, and low-light detection assays (*see* U.S. Patent Nos. 4,376,110 and 4,486,530; *see also Antibodies: A Laboratory Manual* (*supra*)).

A fluorescent antibody test (FA-test) uses a fluorescently labeled antibody able to bind to one of the proteins of the invention. For detection, visual determinations are made by a technician using fluorescence microscopy, yielding a qualitative result. In one embodiment, this assay is used for the examination of tissue samples or histological sections.

In latex bead agglutination assays, antibodies to one or more of the proteins of the present invention are conjugated to latex beads. The antibodies conjugated to the latex beads are then contacted with a sample under conditions permitting the antibodies to bind to desired proteins in the sample, if any. The results are then read visually, yielding a qualitative result. In one embodiment, this format can be used in the field for on-site testing.

Enzyme immunoassays (EIA) include a number of different assays able to utilize the antibodies provided by the present invention. For example, a heterogeneous indirect EIA uses a solid phase coupled with an antibody of the invention and an affinity purified, anti-IgG immunoglobulin preparation. Preferably, the solid

phase is a polystyrene microtiter plate. The antibodies and immunoglobulin preparation are then contacted with the sample under conditions permitting antibody binding, which conditions are well known in the art. The results of such an assay can be read visually, but are preferably read using a spectrophotometer, such as an ELISA plate reader, to yield a quantitative result. An alternative solid phase EIA format includes plastic-coated ferrous metal beads able to be moved during the procedures of the assay by means of a magnet. Yet another alternative is a low-light detection immunoassay format. In this highly sensitive format, the light emission produced by appropriately labeled bound antibodies are quantitated automatically. Preferably, the reaction is performed using microtiter plates.

In an alternative embodiment, a radioactive tracer is substituted for the enzyme mediated detection in an EIA to produce a radioimmunoassay (RIA).

In a capture-antibody sandwich enzyme assay, the desired protein is bound between an antibody attached to a solid phase, preferably a polystyrene microtiter plate, and a labeled antibody. Preferably, the results are measured using a spectrophotometer, such as an ELISA plate reader.

In a sequential assay format, reagents are allowed to incubate with the capture antibody in a step wise fashion. The test sample is first incubated with the capture antibody. Following a wash step, an incubation with the labeled antibody occurs. In a simultaneous assay, the two incubation periods described in the sequential assay are combined. This eliminates one incubation period plus a wash step.

A dipstick/immunostick format is essentially an immunoassay except that the solid phase, instead of being a polystyrene microtiter plate, is a polystyrene paddle or dipstick. Reagents are the same and the format can either be simultaneous or sequential.

In a chromatographic strip test format, a capture antibody and a labeled antibody are dried onto a chromatographic strip, which is typically nitrocellulose or nylon of high porosity bonded to cellulose acetate. The capture antibody is usually spray dried as a line at one end of the strip. At this end there is an absorbent material that is in contact with the strip. At the other end of the strip the labeled antibody is

deposited in a manner that prevents it from being absorbed into the membrane. Usually, the label attached to the antibody is a latex bead or colloidal gold. The assay may be initiated by applying the sample immediately in front of the labeled antibody.

Immunofiltration/immunoconcentration formats combine a large solid phase surface with directional flow of sample/reagents, which concentrates and accelerates the binding of antigen to antibody. In a preferred format, the test sample is preincubated with a labeled antibody then applied to a solid phase such as fiber filters or nitrocellulose membranes or the like. The solid phase can also be precoated with latex or glass beads coated with capture antibody. Detection of analyte is the same as standard immunoassay. The flow of sample/reagents can be modulated by either vacuum or the wicking action of an underlying absorbent material.

A threshold biosensor assay is a sensitive, instrumented assay amenable to screening large numbers of samples at low cost. In one embodiment, such an assay comprises the use of light addressable potentiometric sensors wherein the reaction involves the detection of a pH change due to binding of the desired protein by capture antibodies, bridging antibodies and urease-conjugated antibodies. Upon binding, a pH change is effected that is measurable by translation into electrical potential (μvolts). The assay typically occurs in a very small reaction volume, and is very sensitive. Moreover, the reported detection limit of the assay is 1,000 molecules of urease per minute.

The present invention also provides for probes and primers for detecting *Streptococcus pneumoniae* and *Streptococcus pyogenes*.

In one embodiment of this aspect of the invention, probes are provided that are capable of specifically hybridizing to *S. pneumoniae* and *S. pyogenes* Hsp60 genes DNA or RNA. For purposes of the present invention, probes are "capable of hybridizing" to *S. pneumoniae* and *S. pyogenes* Hsp60 genes DNA or RNA if they hybridize under conditions of high stringency (*see* Sambrook et al. (*supra*)). Preferably, the probe may be utilized to hybridize to suitable nucleotide sequences under highly stringent conditions, such as 6x SSC, 1x Denhardt's solution (Sambrook et al. (*supra*)), 0.1% SDS at 65°C and at least one wash to remove excess probe in the presence of 0.2x

SSC, 1x Denhardt's solution, 0.1% SDS at 65°C. Except as otherwise provided herein, probe sequences are designed to allow hybridization to Streptococcal DNA or RNA sequences, but not to DNA or RNA sequences from other organisms, particularly other bacterial sequences. The probes are used, for example, to hybridize to nucleic acid that has been exposed from a cell in a sample. The hybridized probe is then detected, thereby indicating the presence of the desired cellular nucleic acid. Preferably, the cellular nucleic acid is subjected to an amplification procedure, such as PCR, prior to hybridization.

Probes of the present invention may be composed of either deoxyribonucleic acids (DNA) or ribonucleic acids (RNA), and may be as few as about 12 nucleotides in length, usually about 14 to 18 nucleotides in length, and possibly as large as the entire sequence of the *S. pneumoniae* and *S. pyogenes* Hsp60 genes. Selection of probe size is somewhat dependent upon the use of the probe, and is within the skill of the art.

Suitable probes can be constructed and labeled using techniques that are well known in the art. Shorter probes of, for example, 12 bases can be generated synthetically. Longer probes of about 75 bases to less than 1.5 kb are preferably generated by, for example, PCR amplification in the presence of labeled precursors such as [α-^{32}P]dCTP, digoxigenin-dUTP, or biotin-dATP. Probes of more than 1.5 kb are generally most easily amplified by transfecting a cell with a plasmid containing the relevant probe, growing the transfected cell into large quantities, and purifying the relevant sequence from the transfected cells. (*See* Sambrook et al. (*supra*)).

Probes can be labeled by a variety of markers, including for example, radioactive markers, fluorescent markers, enzymatic markers, and chromogenic markers. The use of ^{32}P is particularly preferred for marking or labeling a particular probe.

It is a feature of this aspect of the invention that the probes can be utilized to detect the presence of *S. pneumoniae* and *S. pyogenes* Hsp60 mRNA or DNA within a sample. However, if the bacteria are present in only a limited number, then it

may be beneficial to amplify the relevant sequence such that it may be more readily detected or obtained.

A variety of methods may be utilized in order to amplify a selected sequence, including, for example, RNA amplification (*see* Lizardi et al., *Bio/Technology 6*:1197-1202, 1988; Kramer et al., *Nature 339*:401-402, 1989; Lomeli et al., *Clinical Chem. 35*(9):1826-1831, 1989; U.S. Patent No. 4,786,600), and DNA amplification utilizing LCR or Polymerase Chain Reaction ("PCR") (*see* U.S. Patent Nos. 4,683,195, 4,683,202, and 4,800,159; *see also* U.S. Patent Nos. 4,876,187 and 5,011,769, which describe an alternative detection/amplification system comprising the use of scissile linkages), or other nucleic acid amplification procedures that are well within the level of ordinary skill in the art. With respect to PCR, for example, the method may be modified as known in the art. PCR may also be used in combination with reverse dot blot hybridization (Iida et al., *FEMS Microbiol. Lett. 114*:167-172, 1993). PCR products may be quantitatively analyzed by incorporation of dUTP (Duplàa et al., *Anal. Biochem. 212*:229-236, 1993), and samples may be filter sampled for PCR-gene probe detection (Bej et al., *Appl. Environ. Microbiol. 57*:3529-3534, 1991).

Within a preferred embodiment, PCR amplification is utilized to detect *S. pneumoniae* and *S. pyogenes* Hsp60 DNA. Briefly a DNA sample is denatured at 95° C in order to generate single-stranded DNA. Specific primers are then annealed to the single-stranded DNA at 37°C to 70°C, depending on the proportion of AT/GC in the primers. The primers are extended at 72°C with *Taq* DNA polymerase in order to generate the opposite strand to the template. These steps constitute one cycle, which may be repeated in order to amplify the selected sequence.

Within an alternative preferred embodiment, LCR amplification is utilized for amplification. LCR primers are synthesized such that the 5' base of the upstream primer is capable of hybridizing to a unique base pair in a desired gene to specifically detect a strain of *Streptococcus* harboring the desired gene.

Within another preferred embodiment, the probes are used in an automated, non-isotopic strategy wherein target nucleic acid sequences are amplified by

PCR, and then desired products are determined by a colorimetric oligonucleotide ligation assay (OLA) (Nickerson et al., *Proc. Natl. Acad. Sci. USA 81*:8923-8927, 1990).

Primers for the amplification of a selected sequence should be selected from sequences that are highly specific and form stable duplexes with the target sequence. The primers should also be non-complementary, especially at the 3' end, should not form dimers with themselves or other primers, and should not form secondary structures or duplexes with other regions of DNA. In general, primers of about 18 to 20 nucleotides are preferred, and can be easily synthesized using techniques well known in the art. PCR products, and other nucleic acid amplification products, may be quantitated using techniques known in the art (Duplàa et al., *Anal. Biochem. 212*:229-236, 1993; Higuchi et al., *Bio/Technology 11*:1026-1030).

Further a biochip array specific for *Streptococcus*, comprised of a substrate to which either oligonucleotides or polypeptides may be bound can be manufactured using the invention disclosed herein in combination with current biochip technologies. U.S. Patent No. 5,445,934. By using such a substrate with oligonucleotides derived from the Streptococcal Hsp60 sequences or antibodies specific for the Streptococcal gene products of this invention, a high throughput screening tool can be created to identify the specific Streptococcal strain in many samples.

E. PHARMACEUTICAL COMPOSITIONS AND METHODS

Another aspect of the present invention provides compositions and methods comprising one or more of the above-described Streptococcal Hsp60 polypeptides or antibodies to Streptococcal Hsp60 in combination with one or more pharmaceutically or physiologically acceptable carriers, adjuvants, binders or diluents. Such compositions can be used to elicit or enhance an immune response in a recipient animal, which is preferably a human being, and preferably elicits or enhances a protective or partially protective immunity against *Streptococcus*, or against an organism associated with an antigen fused to the Streptococcal Hsp60s of the present invention.

Preferably, such carriers, adjuvants, binders or diluents are nontoxic to recipients at the dosages and concentrations employed. Ordinarily, the preparation of such compositions entails combining the isolated Streptococcal Hsp60 polypeptide with buffers, antioxidants such as ascorbic acid, low molecular weight (less than about 10 residues) polypeptides, proteins, amino acids, carbohydrates including glucose, sucrose or dextrins, chelating agents such as EDTA, glutathione and other stabilizers and excipients. Neutral buffered saline or saline mixed with nonspecific serum albumin are exemplary appropriate diluents. Examples of adjuvants include alum or aluminum hydroxide for humans.

It will be evident in light of the present specification to those in the art that the amount and frequency of administration can be optimized in clinical trials, and will depend upon such factors as the disease or disorder to be treated, the degree of immune inducement, enhancement, or protection required, and many other factors.

In one embodiment, the composition is administered orally, and the purified Streptococcal Hsp60 is taken up by cells, such as cells located in the lumen of the gut. Alternatively, the Streptococcal Hsp60 composition can be parenterally administrated via the subcutaneous route, or via other parenteral routes. Other routes include buccal/sublingual, rectal, nasal, topical (such as transdermal and ophthalmic), vaginal, pulmonary, intraarterial, intramuscular, intraperitoneal, intraocular, intranasal or intravenous, or indirectly. The Streptococcal Hsp60 compositions of the present invention can be prepared and administered as a liquid solution, or prepared as a solid form (*e.g.*, lyophilized) which can be administered in solid form or resuspended in a solution in conjunction with administration.

Depending upon the application, quantities of injected Streptococcal Hsp60 in the composition will vary generally from about 0.1 μg to 1000 mg, typically from about 1 μg to 100 mg, preferably from about 10 μg to 10 mg, and preferably from about 100 μg to 1 mg, in combination with the physiologically acceptable carrier, binder or diluent. Booster immunizations can be given from 2-6 weeks later.

The pharmaceutical compositions of the present invention may be placed within containers, along with packaging material, preferably consumer-acceptable,

which provides instructions regarding the use of such pharmaceutical compositions, to provide kits suitable for use within the present invention. Generally, such instructions will include a tangible expression describing the reagent concentration, as well as within certain embodiments, relative amounts of excipient ingredients or diluents (*e.g.*, water, saline or PBS) which may be necessary to reconstitute the pharmaceutical composition.

The Hsp gene products of this invention may also be used as immunological carriers in conjugate vaccines. Hsps are beneficial carriers of antigens because, unlike other carriers, they do not have an immunosuppressive effect. See Barrios et al., *Eur. J. Immunol. 22*:1365-1372, 1992; Suzue and Young, in *Stress-Inducible Cellular Responses 77*:451-465, 1996 (edited by U. Feige et al.). Such carriers may be used to elicit an increased immune response to the conjugated molecule. The Streptococcal Hsp gene products of this invention may therefore be used as carriers (in conjugates or fusion proteins).

An additional aspect of the present invention is the use of the Streptococcal Hsp60 genes and gene products to treat and/or prevent tumors. The methods comprise administering to an individual having cancer a composition comprising a purified Streptococcal Hsp60 gene product as discussed herein in an amount effective to elicit and/or enhance the immune response of an individual against the cancer. The present invention also provides a method of immunizing an individual against cancer, or of providing at least a partially effective immunoprotective response in such an individual, the method comprising administering to the individual a composition comprising a purified Streptococcal Hsp60 as discussed herein in an amount effective to immunize the individual.

Preferably, the treatment of cancer comprises the use of highly purified Streptococcal Hsp60 gene products that are substantially free of endotoxins and methods and compositions related to the same. Such highly purified proteins are particularly advantageous, for example, for the treatment of human cancers because they do not incur the adverse side effects associated with such endotoxins. In particular, the compositions are capable of inducing an immune response against a cancer existing

within an individual, which includes both eliciting the immune response or enhancing the immune response against the cancer. For example, the cancer to be treated may be an endothelial cell cancer, such as a sarcoma and/or breast, ovarian, prostate, lung, pancreas and liver cancers. The present invention also provides compositions that are capable of providing either partially or fully protective immune responses by immunization against cancers that are not yet present within an individual.

A further aspect of the present invention is protection from a variety of bacterial diseases by either immunization with the Hsp60 gene products of the present invention or by using gene transfer techniques to deliver a vector containing Streptococcal Hsp60 genes or fragments thereof to be expressed within the cells of the animal. The compositions and methods of the present invention can also provide for cancer prevention.

The compositions and methodologies described herein are suitable for a variety of uses. To this end, the following examples are presented for purposes of illustration, not limitation.

EXAMPLES

EXAMPLE 1

ISOLATION OF GENES FOR STREPTOCOCCUS PNEUMONIAE AND STREPTOCOCCUS PYOGENES HSPS

Genomic DNA from *Streptococcus pneumoniae* (ATCC6314) and *Streptococcus pyogenes* (ATCC12344), prepared by a routine method, wasobtained from Dr. Lee Weber, University of Nevada at Reno.

Hsp60 DNA sequences were isolated by use of the polymerase chain reaction. Primers were designed based on N- and C-terminal homology of known Hsp60 sequences from other organisms. DNA amplifications of Streptococcal DNA were carried out using Taq polymerase (Perkin-Elmer). About 20 different primer pairs were tested using different reaction conditions. One pair (pair 1) was identified that was

capable of amplifying Hsp60-1 genes, and a second (pair 2) that permitted amplification of Hsp60-2 sequences. Reaction mixtures capable of amplifying Hsp60 sequences contained, in a total volume of 50 µl, 0.5 µg of genomic DNA, 50 pmoles of each of a pair of degenerate primers, 500 µM each of dNTPs, 1xPCR buffer (Perkin-Elmer), 2 mM $MgSO_4$, and 1.25 units of Taq polymerase (Perkin-Elmer). The following two pairs of degenerate primers were employed successfully:

Pair 1:

forward primer #1F: 5'-CATATGGCNGCNAAAGAYGTAAAA-3' (SEQ ID NO:35)

reverse primer #1R: 5'-TGATCACATCATNCCNCCCATNCC-3' (SEQ ID NO:36)

Pair 2:

forward primer #2F: 5'-CATATGGCAAAAGAAATHAARTTY-3' (SEQ ID NO:37)

reverse primer #2R: 5'-TGATCANCCNCCCATNCCNCCCAT-3' (SEQ ID NO:38)

In the above sequences, N refers to A, C, G or T, and H to A, C or T (not G).

Reactions were cycled 35 times at 94°C for 1 minute, 50°C for 2 minutes and 72°C for 2 minutes. PCR products were electrophoresed on 0.6% low-melting point agarose gels (Gibco-BRL) along with molecular weight markers. After staining with ethidium bromide, DNA fragments were visualized under low-intensity, long-wavelength UV illumination, and fragments of about 1.6 kbp were excised. DNA was isolated from gel slices by phenol extraction and ethanol precipitation (Maniatis et al.). Purified fragments were ligated to pCRII TA cloning vector (Invitrogen), and ligation mixtures were used to transform *E. coli* strain DH5a. (Competent cells obtained from Life Technologies.) Recombinant plasmids were isolated from kanamycin-resistant colonies by a standard alkaline lysis method, and the presence in plasmids of DNA inserts was verified by digestion with EcoRI digestion followed by agarose gel electrophoresis and visualization of digestion products by staining with ethidium bromide.

EXAMPLE 2

NUCLEOTIDE SEQUENCE ANALYSIS OF STREPTOCOCCAL HSP60

Inserts present in recombinant pCRII-based clones were sequenced using a CircumVent sequencing kit (New England Biolabs), ^{35}S-dATP and primers listed below. Multiple clones containing particular Streptococcal *Hsp60* genes were sequenced: sequences were obtained from five clones, derived from three independent PCR reactions, of the *Streptococcus pneumoniae Hsp60-1* gene, two clones, derived from single PCR reactions, of the *Streptococcus pneumoniae Hsp60-2* gene, four clones, derived from three independent PCR reactions, of the *Streptococcus pyogenes Hsp60-1* gene, and two clones and a portion of a third clone, derived from single PCR reaction, of the *Streptococcus pyogenes Hsp60-2* gene. Sequencing reactions were fractionated on denaturing 6% polyacrylamide-8M urea gels (60 cm length), and the gels were dried and exposed for autoradiography. Autoradiographs were read manually, and sequence data were assembled and compared to other known *Hsp60* genes using DNA Strider software (CEA, France).

Sequencing primers used:

M13F : 5'-GTAAAACGACGGCCAG-3' (SEQ ID NO:39)
M13R : 5'-CAGGAAACAGCTATGAC-3' (SEQ ID NO:40)
W178 : 5'-CCAACCATCACGAAAGA-3' (SEQ ID NO:41)
W179 : 5'-ACGGGTCACTTTGGTTG-3' (SEQ ID NO:42)
W189 : 5'-TTACTAATGACGGGGTA-3' (SEQ ID NO:43)
W190 : 5'-TTACCAATGACGGTGTG-3' (SEQ ID NO:44)
W191 : 5'-ACAGGGTCAATGATTCC-3' (SEQ ID NO:45)
W192 : 5'-ACTGGATCAATGATACC-3' (SEQ ID NO:46)
W195 : 5'-CCGTACCGTGCTCTGAC-3' (SEQ ID NO:47)
W196 : 5'-ACCACGTTTCAGATCCA-3' (SEQ ID NO:48)

W197 : 5'-GACAGTTTCGCGGCAAC-3' (SEQ ID NO:49)
W198 : 5'-CTCAGAACGAAGATCAG-3' (SEQ ID NO:50)
W200 : 5'-GGTATGCAGTTCGACCG-3' (SEQ ID NO:51)
W201 : 5'-CCGTGTTGGTCAAATCC-3' (SEQ ID NO:52)
W202 : 5'-GGTAACTACGGTTACAA-3' (SEQ ID NO:53)
W203 : 5'-GAGGCCACTTCTTTCAC-3' (SEQ ID NO:54)
W204 : 5'-GGCTTCCAGCACTGGCA-3' (SEQ ID NO:55)
W205 : 5'-AACTTCAGTCGCAGCAC-3' (SEQ ID NO:56)
W206 : 5'-CCTTGAAAGCCATTGCT-3' (SEQ ID NO:57)
W207 : 5'-GCTACACGTGCAGCCGT-3' (SEQ ID NO:58)
W208 : 5'-GCTGCAACAGGTGAGTG-3' (SEQ ID NO:59)
W209 : 5'-TCATGAACAATGGCTTG-3' (SEQ ID NO:60)
W210 : 5'-ACGAAGCACAATGTTAC-3' (SEQ ID NO:61)
W211 : 5'-ATCACTAAAGATGGTGT-3' (SEQ ID NO:62)
W214 : 5'-GCAGTTGCCGCAGCAGT-3' (SEQ ID NO:63)
W215 : 5'-GCTACTCGTGCAGCTGT-3' (SEQ ID NO:64)
W216 : 5'-GTTCTCCGTGCTTTGGA-3' (SEQ ID NO:65)
W217 : 5'-GCACCTGCTGTGACGTT-3' (SEQ ID NO:66)
W218 : 5'-TCTTCGATGGTGATGAC-3' (SEQ ID NO:67)
W219 : 5'-GGCAAGAGCTGTTCCGC-3' (SEQ ID NO:68)
W220 : 5'-CTGAGCCAGTACGGTTG-3' (SEQ ID NO:69)
W221 : 5'-GTACTGCAGAGCGGAAC-3' (SEQ ID NO:70)
W224 : 5'-ACCGTCTTCAACGGTGA-3' (SEQ ID NO:71)
W225 : 5'-GTTATCATTGCTGAAGA-3' (SEQ ID NO:72)
W226 : 5'-ACGGTACCGCCGGTCAG-3' (SEQ ID NO:73)
W227 : 5'-CTGGGCCAGGCTAAACG-3' (SEQ ID NO:74)
W228 : 5'-CGACTGAAGTTGAAATG-3' (SEQ ID NO:75)
W229 : 5'-GCTGTTGAAGAACTGAA-3' (SEQ ID NO:76)
W230 : 5'-GTCTTCAACGGTGATCA-3' (SEQ ID NO:77)
W232 : 5'-TCTTCTACCGCAGCACG-3' (SEQ ID NO:78)

W233 : 5'-CTCTTGATTATTGCGGA-3' (SEQ ID NO:79)
W234 : 5'-TTGTTCAAAACAAGAGT-3' (SEQ ID NO:80)
W235 : 5'-CGATTATTGTAGAAGGT-3' (SEQ ID NO:81)
W236 : 5'-CTTGATAACCGCAACAC-3' (SEQ ID NO:82)
W237 : 5'-TCCAAAGCACGGAGAAC-3' (SEQ ID NO:83)
W238 : 5'-GTGTCAAACATCCAAGA-3' (SEQ ID NO:84)
W239 : 5'-TCTTCGATGGTAATCAC-3' (SEQ ID NO:85)
W240 : 5'-GCAATAATGAGTAATGG-3' (SEQ ID NO:86)
W241 : 5'-ACAGTAATTGTTGAAGG-3' (SEQ ID NO:87)
W242 : 5'-CAGTGCAATACGGTTAG-3' (SEQ ID NO:88)
W243 : 5'-AGCTTCCAGAACCGGCA-3' (SEQ ID NO:89)
W244 : 5'-CTGATCATCGCTGAAGA-3' (SEQ ID NO:90)
W245 : 5'-ACGGTTATTGTAGAAG-3' (SEQ ID NO:91)

The sequencing strategy for each of the Hsp60 genes is set forth in Figure 5. The nucleotide sequences of the *Streptococcus pneumoniae Hsp60-1* gene (referred to as P60-1), the *Streptococcus pneumoniae Hsp60-2* gene (P60-2), the *Streptococcus pyogenes Hsp60-1* gene (Y60-1) and the *Streptococcus pyogenes Hsp60-2* gene (Y60-2), and the corresponding deduced amino acid sequences, are set forth in Figures 1-4 (SEQ ID NOS:1-8).

Comparisons of Streptococcal Hsp60 proteins and mycobacterial Hsp65 and GroEL proteins were determined using the MegAlign module of a DNA Star software package (DNASTAR, Inc.), and sequence similarities to Genbank-listed genes and proteins were uncovered using the BLAST algorithm (National Center for Biotechnology Information, NIH, Bethesda, MD). One comparison of such sequences is set forth in Figure 10.

EXAMPLE 3

EXPRESSION OF RECOMBINANT STREPTOCOCCAL HSP60

Inserts *(Hsp60 genes)* were excised from recombinant pCRII-based plasmids with restriction enzymes NdeI and EcoRI. NdeI cut inside forward PCR primers #1F or #2F, and EcoRI cut a short distance downstream from reverse PCR primers #1R or #2R in the polylinker region of vector PCRII. DNA fragments including *Hsp60* gene sequences were fractionated on low-melting-point agarose gels, purified from the gels and ligated into NdeI/EcoRI double-digested pET28a(+) vector DNA (Novagen). Ligation reactions were used to transform competent *Escherichia coli* DH5a cells, and transformants were selected on Luria Broth plates containing 30μg/ml of kanamycin D. DNA was isolated from single colonies using a standard alkaline lysis method, and the presence of correct inserts verified by digestion with NdeI and EcoRI and agarose gel electrophoresis. The resulting expression plasmids contained either a *Streptococcus pneumoniae Hsp60-1* gene (referred to as pETP60-1), a *Streptococcus pneumoniae Hsp60-2* gene (pETP60-2), *a Streptococcus pyogenes Hsp60-1* gene (pETY60-1) or a *Streptococcus pyogenes Hsp60-2* gene (pETY60-2). Schematic maps of the expression plasmids are shown in Figures 6-9.

To test whether they were capable of expressing the inserted Streptococcal *Hsp60* genes, the expression plasmids were introduced into *Escherichia coli* strain BL21(DE3) by electroporation, and transformant colonies were selected on kanamycin-containing plates as before. Cultures of one ml were inoculated with single colonies, and transformants were grown at 37°C, until the cultures were turbid. After removing an aliquot for analysis of proteins prior to induction of recombinant genes (uninduced cultures), isopropyl-thio-galactopyranoside (IPTG) was added to 1mM, and cultures were incubated for an additional one or two hours (induced cultures). Aliquots of 100μl of induced and uninduced cultures were centrifuged at 12,000 x g for 30 seconds. Bacterial pellets were lysed in 100μl of SDS-PAGE loading buffer and boiled for 3 minutes. Aliquots of 10μl of lysates were analyzed by 10% SDS-PAGE. Recombinant Streptococcal Hsp60 proteins were detectable after Coomasie blue

staining as prominent bands migrating with an apparent molecular weight of about 60kDa, which bands were present in induced but not in uninduced samples.

EXAMPLE 4

PURIFICATION OF RECOMBINANT STREPTOCOCCAL HSP60

Bacteria containing recombinant Streptococcal *Hsp60* expression plasmids were grown in 2xYT medium (20 g Tryptone, 10 g yeast extract, 10 g NaCl per liter) supplemented with 30μg/ml of kanamycin D at 37°C to an optical density at 600 nm of 0.5-0.8 and then induced with 0.5 mM IPTG for 3 hours. Cultures were then chilled on ice, and bacteria collected by centrifugation at 7,000 x g for 5 min (at 4°C). Bacterial pellets were frozen at -80°C.

Frozen bacterial pellet was crushed, transferred to a blender and homogenized in 200ml of buffer A (6 M guanidinium hydrochloride, 50 mM Tris-HCl pH 7.5, 0.5 mM beta-mercaptoethanol).

Lysate was cleared by centrifugation at 10,000 x g for 15 min (at 4°C). The supernatant solution was mixed overnight at room temperature with approximately 100ml of slurry containing 50ml of Ni-Sepharose (Chelating Sepharose, Pharmacia) equilibrated in buffer A. The resin was then washed on filter paper with approximately 200 ml buffer A, resuspended in small volume of the same buffer and gravity-packed into glass chromatography column (Pharmacia).

The column was washed with 20 ml of buffer A with 1% Triton X-100. The column was further washed with a 6 - 0 M guanidinium hydrochloride / 0 - 1 M NaCl gradient in 50 mM Tris-HCl pH 7.5, 0.5 mM beta-mercaptoethanol (200ml), then with 200 ml of 50 mM Tris-HCl pH 7.5, 1 M NaCl, 0.5 mM beta-mercaptoethanol, and finally with 200 ml of 50 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl pH 7.5, 1.025 M NaCl, 0.5 mM beta-mercaptoethanol. Then column was developed with a 200ml-gradient from 5% to 100% of buffer composed of 1 M imidazole, 0.5 M NaCl, 50 mM Tris-HCl pH 7.5, 0.5 mM beta-mercaptoethanol in 1M NaCl, 50 mM Tris-HCl pH 7.5,

0.5 mM beta-mercaptoethanol. Fractions of 9ml were collected. The flow rate was 4-5 ml/min, and chromatography was monitored by absorbance at 280 nm.

Fractions containing the highest concentrations of recombinant protein were identified by 10% SDS-PAGE as before, pooled (usually 5-6 fractions) into a dialysis bag (12 kDa cutoff). Protein solution (approximately 50 ml) was then dialysed in the cold against three changes of 3 liters of Dulbeccos' phosphate-buffered saline (2.7 mM KH_2PO_4, 4.3 mM Na_2HPO_4, 2.7 mM KCl, 0.137 M NaCl). Dialysed protein was aliquoted and stored at -80°C. Usually 200-400 mg of recombinant protein were obtained (estimated by protein assay according to Lowry).

EXAMPLE 5

CHARACTERIZATION OF PURIFIED, RECOMBINANT HSP60

To unambiguously identify recombinant proteins as Streptococcal Hsp60, purified recombinant proteins were subjected to

N- and C-terminal sequencing (conducted by the Protein Chemistry Facility, W. Alton Jones Cell Science Center, Lake Placid, NY). These determinations revealed that purified recombinant proteins had the C- and N-terminal sequences predicted from the deducted amino acid sequences of SEQ ID NOS:5-8 (except for the N-terminal methionine that is typically processed away in *E. coli* bacteria).

EXAMPLE 6

REACTIVITY OF RECOMBINANT STREPTOCOCCAL HSP60 WITH KNOWN ANTI-HSP60 MONOCLONAL ANTIBODIES

Purified recombinant Streptococcal Hsp60 proteins were analyzed for reactivity with the following commercially available antibodies:

A) Rabbit polyclonal antibody SPA-804 (StressGen Biotechnologies) which was raised against *Synechococcus sp.* Hsp60. The antibody recognizes Hsp60 from a wide range of prokaryotes and eukaryotes

including cyanobacteria, *Escherichia coli*, and primate, murine, hamster, and rat cell lines.

B) Murine monoclonal antibody SPA-807 (StressGen Biotechnologies) which was raised against human Hsp60. Its epitope is located between residues 383-419 of that protein. The antibody also cross-reacts with Hsp60 from various other species including primates, rabbit, mouse, rat, hamster, *Borrelia sp.*, *Escherichia coli*, *Streptococcus pyogenes*, *Yersinia enterocolitica*, *Salmonella typhimurium*, *Treponema hyodysenteriae*, *Treponema innocense*, *Trichinella spiralis*, yeast, and spinach chloroplasts.

C) Murine monoclonal antibody SPA-870 (StressGen Biotechnologies) which was raised against *Escherichia coli* GroEL. The antibody does not recognize eukaryotic Hsp60 proteins.

D) Murine polyclonal antibody which was raised against *Mycobacterium tuberculosis* BCG Hsp60 (StressGen Biotechnologies). The antibody does not cross-react with *Escherichia coli* groEL or eukaryotic Hsp60.

E) Murine monoclonal antibody recognizing recombinant histidine tag (Qiagen).

Samples containing 0.1μg, 0.5μg or 1μg of recombinant protein were fractionated on 10% SDS-PAGE, and proteins were electroblotted onto nitrocellulose. Blots for analysis with antibodies SPA-804, SPA-807, SPA-870, and anti-BCG Hsp60 were blocked with 5% skim milk in PBS containing 0.05% Tween 20 overnight at room temperature. Blots were then incubated for one hour in the same buffer containing primary antibody (at a 1:1000 dilution except for anti-BCG Hsp60 antibody which was used at a 1:500 dilution). Blots were washed 3 times (10 min each) with PBS with 0.05% Tween 20 and incubated for an additional hour in PBS with 5% skim milk, 0.05% Tween 20 and goat anti-rabbit IgG - alkaline phosphatase (AP) conjugate (Sigma) or goat-anti-murine IgG - alkaline phosphatase (AP) conjugate (Sigma) (at

1:1000 dilutions), respectively. After 3 washes in PBS with 0.05% Tween 20 as before, blots were soaked in alkaline phosphatase reaction buffer (100 mM Tris-HCl (pH 9.5), 150 mM NaCl, 10 mM $MgCl_2$) and then developed in 0.05% nitroblue tetrazolium (NBT), 0.05% 5-bromo-4-chloro-3-indolyl phosphate (BCIP) in the same buffer, until signals were clearly visible (approximately 15 minutes).

A similar procedure was followed for anti-histidine tag antibody, except that blocking was in 3% bovine serum albumin in TBS (10 mM Tris-HCl, pH 7.5, 150 mM NaCl). Primary and secondary antibodies were diluted in TBS alone, and incubation with primary antibody (1:500 dilution) was for two hours. Washes were performed as follows: blots were first washed twice for 10 min in TBS containing 0.05% Tween 20 and 0.2% Triton X-100, and once for 10 min in TBS.

Recombinant histidine-tagged Hsp60 proteins were purified from overexpressing *E. coli* cells and probed on Western blot with polyclonal antibodies SPA-804 and anti-BCG Hsp60 as well as monoclonal antibodies SPA-870, SPA-807, and anti-histidine tag antibody. As is shown in Table 1, SPA-804 recognized all four Streptococcal Hsp60 proteins. In contrast, SPA-807 failed to crossreact with *Streptococcus pneumoniae* Hsp60-2, SPA-870 was unable to react with any Streptococcal Hsp60-2 protein, and anti-BCG Hsp60 failed to crossreact with any Streptococcal Hsp60. As predicted, anti-His tag antibody reacted with all recombinant proteins which had been expressed as His-tagged proteins. Positive reactivity is indicated as "+" while lack of it is marked with "-".

TABLE 1

RECOGNITION OF STREPTOCOCCAL HSP60 PROTEINS BY ANTI-HSP60 ANTIBODIES

Antibody	*S. pneumoniae* Hsp60-1	*S. pneumoniae* Hsp60-2	*S. pyogenes* Hsp60-1	*S. pyogenes* Hsp60-2
SPA-804	+	+	+	+
SPA-807	+	-	+	+
SPA-870	+	-	+	-
anti-BCG60	-	-	-	-
anti-His tag	+	+	+	+

These data demonstrate that Streptococcal Hsp60 are antigenically distinct from Hsp60 of other organisms. They also show that Streptococcal Hsp60-1 and Hsp60-2 can be distinguished. And, they provide evidence that related Hsp60s from two different Streptococcal species can be recognized differentially by an antibody.

EXAMPLE 7

PREPARATION AND IDENTIFICATION OF PEPTIDE FRAGMENTS OF RECOMBINANT STREPTOCOCCAL HSP60

Purified recombinant proteins (50 mg at 1 mg/ml) were digested with 2.5 mg of Lys-C endopeptidase (Boehringer Mannheim) for 1 hour at 37°C. Digestion reactions were fractionated by capillary electrophoresis (3D-HPCE instrument, Hewlett-Packard). Reactions were run at 15 kV through a 75 u bare fused silica capillary in 50 mM dibasic sodium phosphate (pH 7.47). Alternatively, reverse phase chromatography (1100 Series HPLC instrument, Hewlett-Packard) was carried out on a Hamilton PRP-1 5 m column developed in a 0-60% acetonitrile gradient in water in the presence of 0.1% trifluoroacetic acid. Individual RP-HPLC-separated peptides of Hsp60 proteins were identified by mass spectroscopy by Hewlett-Packard Laboratories, Palo Alto, California. RP-HPLC chromatograms of digests of recombinant Streptoccocal Hsp60s are shown in Figure 11.

EXAMPLE 8

IDENTIFICATION OF ENDOGENOUS STREPTOCOCCAL HSP60

Total protein extracts from *Streptococcus pneumoniae* (ATCC6314) and *Streptococcus pyogenes* (ATCC12344) were obtained from Dr. Lee Weber (University of Nevada, Reno). Equivalent amounts of both extracts (equalized based on intensity of staining of protein bands in SDS-PAGE gels) were fractionated by 10% SDS-PAGE alongside 50 ng of purified BGC Hsp60 (StressGen Biotechnologies). After

electrotransfer onto nitrocellulose, filters were blocked, probed with antibody SPA-804, and antibody signals detected as described in Example 6.

Other, similarly prepared filters were incubated with a 1:3000 dilutions of antibodies SPA-807 or SPA-870 for one hour. Blots were rinsed twice with water, washed 3 times (5 min each) with PBS containing 0.05% Tween 20 and then incubated for an additional hour in PBS containing 5% skim milk, 0.05% Tween 20 and a 1:3000 dilution of goat anti-rabbit IgG - horseradish peroxidase (HRP) conjugate (Sigma). Subsequently, filters were rinsed with water, washed with PBS containing 0.05% Tween 20 as before, equilibrated in ECL substrate mixture (Amersham), wrapped in plastic wrap and exposed to X-ray film for between 15 seconds and 20 minutes.

The results from these experiments are summarized in Table 2. Antibody SPA-804 reacted strongly with both Streptococcal extracts. In contrast, antibody SPA-807 reacted weakly with extract from *Streptococcus pneumoniae* but strongly with extract from *Streptococcus pyogenes*. Finally, antibody SPA-870 reacted weakly with both Streptococcal extracts. Based on the antibody specificity determined in Example 6 (Table 1), it is concluded that Hsp60-2 is abundant in Streptococcal cells, whereas Hsp60-1 is only expressed at low levels. Presumably, Hsp60-1 is the more highly stress-inducible Hsp60 protein.

TABLE 2

REACTIVITY OF SELECTED ANTI-HSP60 ANTIBODIES WITH PROTEIN EXTRACTS FROM *S. PNEUMONIAE AND S. PYOGENES*

Antibody	BCG60 control (50 ng)	*S. pneumoniae* extract	*S. pyogenes* extract
SPA-804	+++	+++	+++
SPA-807	+++	+	+++
SPA-870	-	+	+ a

a: SPA870 detected protein with mobility different from predominant heavy band visualized in that extract with SPA-807. However, its mobility was close to the band detected in *S. pneumoniae* extract with both SPA-870 and SPA-807 antibodies.

The amount of the utilized extracts was normalized by comparing Coomassie stained gels containing serial dilutions.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

1. An isolated nucleic acid molecule encoding a *Streptococcus pneumoniae* Hsp60.

2. An isolated nucleic acid molecule encoding a *Streptococcus pyogenes* Hsp60.

3. An isolated nucleotide molecule selected from the group consisting of:

(a) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:1 from nucleotides 15-1652;

(b) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:3 from nucleotides 15-1640;

(c) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:5 from nucleotides 15-1649;

(d) an isolated nucleic acid molecule comprising the sequence of SEQ ID NO:7 from nucleotides 15-1652;

(e) an isolated nucleic acid molecule complementary to any one of the nucleotides of SEQ ID NOS:1, 3, 5 or 7 set forth in (a) through (d), respectively; and

(f) an isolated nucleic acid molecule that hybridizes under conditions of high stringency to the nucleic acid molecules of any one of (a) through (e).

4. An isolated nucleic acid molecule that specifically hybridizes to the nucleic acid molecule of any one of SEQ ID NO:1 from nucleotides 15-1652, SEQ ID NO:3 from nucleotides 15-1640, SEQ ID NO:5 from nucleotides 15-1649, or SEQ ID NO:7 from nucleotides 15-1652 or a complement thereof under conditions of high stringency.

5. An isolated nucleic acid molecule comprising a nucleotide sequence that is identical to a segment comprising at least 25% of contiguous nucleotide bases of any one of SEQ ID NO:1 from nucleotides 15-1652, SEQ ID NO:3 from nucleotides 15-1640,

SEQ ID NO:5 from nucleotides 15-1649, or SEQ ID NO:7 from nucleotides 15-1652 or a complement thereof.

6. An isolated nucleic acid molecule encoding Hsp60 comprising a nucleic acid sequence that encodes a polypeptide comprising any one of SEQ ID NOS: 2, 4, 6 or 8 or a variant Hsp60 that is at least 95% homologous to a polypeptide according to any one of SEQ ID NOS: 2, 4, 6 or 8.

7. An isolated nucleic acid molecule according to claim 3, encoding a polypeptide that is able to be selectively bound by an antibody specific for a *Streptococcus pneumoniae* Hsp60 or a *Streptococcus pyogenes* Hsp60.

8. An isolated nucleic acid molecule encoding at least 8 amino acids of a Streptococcal Hsp60 polypeptide selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-541 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, wherein the encoded Streptococcal Hsp60 polypeptide is able to bind to a major histocompatibility complex.

9. An isolated *Streptococcus pneumoniae* Hsp60 polypeptide.

10. An isolated *Streptococcus pyogenes* Hsp60 polypeptide.

11. An isolated Hsp60 polypeptide comprising the amino acid sequence of any one of a Streptococcal Hsp60 polypeptide selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-541 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, or variants thereof, wherein the polypeptide is able to be selectively bound by an antibody specific for either a *Streptococcus pneumoniae* Hsp60 and/or *Streptococcus pyogenes* Hsp60.

12. The isolated Hsp60 polypeptide according to any one of claims 9-11, wherein the Hsp60 polypeptide is fused to an additional polypeptide to create a fusion protein.

13. An isolated Hsp60 polypeptide comprising at least 8 amino acids selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-541 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, wherein the Hsp60 polypeptide is capable of binding to a major histocompatibility complex and eliciting or enhancing an immune response to *Streptococcus* in a human being.

14. The isolated Hsp60 polypeptide according to claim 11 wherein the polypeptide is derived from proteolytic cleavage.

15. The isolated Hsp60 polypeptide according to claim 11 wherein the polypeptide is derived from chemical synthesis.

16. The isolated Hsp60 according to claim 11 wherein the Hsp60 is an expression product of a transformed host cell containing a nucleic acid molecule encoding the Hsp60 or portion thereof.

17. The isolated Hsp60 polypeptide according to claim 11 wherein the polypeptide comprises greater than 95% homology to any one of a Streptococcal Hsp60 polypeptide selected from amino acid residues 1-545 of SEQ ID NO:2, amino acid residues 1-5410 of SEQ ID NO:4, amino acid residues 1-544 of SEQ ID NO:6, and amino acid residues 1-545 of SEQ ID NO:8, and wherein the Hsp60 polypeptide is able to be selectively bound by an antibody specific for either a *Streptococcus pneumoniae* Hsp60 or *Streptococcus pyogenes* Hsp60 or both.

18. An isolated polypeptide wherein the polypeptide is an expression product of a transformed host cell containing the nucleic acid molecule of any one of claims 1-8.

19. A vector comprising an isolated nucleic acid molecule according to any one of claims 1-8.

20. The vector according to claim 19 wherein the vector is an expression vector comprising a promoter in operative linkage with the isolated nucleic acid molecule encoding the Hsp60 or portion thereof.

21. The vector according to claim 20, further comprising a selectable or identifiable marker.

22. The vector according claim 20 wherein the promoter is a constitutive or an inducible promoter.

23. A host cell containing a vector according to claim 19.

24. The host cell according to claim 24 wherein the host cell is selected from the group consisting of a bacterial cell, a mammalian cell, a yeast cell and an insect cell.

25. A composition comprising an Hsp60 polypeptide of any one of claims 9-16 in combination with a pharmaceutically acceptable carrier or diluent.

26. The composition according to claim 25 wherein the composition is suitable for systemic administration.

27. The composition according to claim 25 wherein the composition is suitable for oral administration.

28. The composition according to claim 25 wherein the composition is suitable for parenteral administration.

29. A method for eliciting or enhancing an immune response in a mammal against *Streptococcus*, comprising administering to the mammal an effective amount of an Hsp60 polypeptide according to any one of claims 9-16 in combination with a pharmaceutically acceptable carrier or diluent.

30. A method for eliciting or enhancing an immune response in a mammal against a target antigen comprising administering to the mammal the target antigen joined to an Hsp60 polypeptide according to any one of claims 9-16 in combination with a pharmaceutically acceptable carrier or diluent.

31. A composition comprising an isolated nucleic acid molecule of any one of claims 1-8 wherein the isolated nucleic acid molecule encodes a polypeptide having at least one amino acid difference from a corresponding polypeptide of an Hsp60 protein from an organism other than *Streptococcus*.

S. pneumoniae Hsp60-1 gene (SEQ ID NOS:1 and 2)

```
GAATTCGGCT TCAT ATG GCG GCT AAA GAC GTA AAA TTC GGT AAC GAC GCT CGT GTG AAA ATG CTG CGC GGC GTA AAC   77
                Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met Leu Arg Gly Val Asn   21

GTA CTG GCA GAT GCA GTG AAA GTT ACC CTC GGC CCA AAA GGC CGT AAC GTA GTT CTG GAT AAA TCT TTC GGT GCA  152
Val Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala   46

CCG ACC ATC ACT AAA GAT GGT GTT TCC GTA GCA CGT GAA ATC GAA CTG GAA GAC AAG TTC GAA AAC ATG GGT GCG  227
Pro Thr Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile Glu Leu Glu Asp Lys Phe Glu Asn Met Gly Ala   71

CAG ATG GTG AAA GAA GTT GCC TCT AAA GCG AAC GAC GCT GCA GGT GAC GGT ACC ACC ACC GCA ACC GTA CTG GCT  302
Gln Met Val Lys Glu Val Ala Ser Lys Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala   96

CAG TCC ATC ATC ACT GAA GGC CTG AAA GCC GTT GCT GCG GGC ATG AAC CCG ATG GAT CTG AAA CGT GGT ATC GAC  377
Gln Ser Ile Ile Thr Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn Pro Met Asp Leu Lys Arg Gly Ile Asp  121

AAA GCT GTC GCT GCT GCT GTT GAA GAA CTG AAA GCA CTG TCC GTA CCG TGC TCC GAC TCT AAA GCT ATT GCT CAG  452
Lys Ala Val Ala Ala Ala Val Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp Ser Lys Ala Ile Ala Gln  146

GTT GGT ACC ATC TCC GCT AAC TCC GAC GAA ACC GTA GGT AAA CTG ATC GCT GAA GCG ATG GAC AAA GTC GGT AAA  527
Val Gly Thr Ile Ser Ala Asn Ser Asp Glu Thr Val Gly Lys Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys  171

GAA GGC GTG ATC ACC GTT GAA GAC GGT ACC GGT CTG CAG GAC GAA CTG GAC GTG GTT GAA GGT ATG CAG TTC GAC  602
Glu Gly Val Ile Thr Val Glu Asp Gly Thr Gly Leu Gln Asp Glu Leu Asp Val Val Glu Gly Met Gln Phe Asp  196

CGT GGC TAC CTG TCT CCT TAC TTC ATC AAC AAG CCG GAA ACT GGC GCA GTA GAA TTG GAA AGC CCG TTC ATC CTG  677
Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Lys Pro Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu  221

CTG GCT GAC AAG AAA ATC TCC AAC ATC CGC GAA ATG CTG CCG GTT CTG GAA GCT GTA GCG AAA GCA GGC AAA CCG  752
Leu Ala Asp Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu Glu Ala Val Ala Lys Ala Gly Lys Pro  246

CTG CTG ATC ATC GCT GAA GAT GTT GAA GGC GAA GCG CTG GCA ACT CTG GTT GTT AAC ACC ATG CGC GGT ATC GTA  827
Leu Leu Ile Ile Ala Glu Asp Val Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile Val  271

AAA GTC GCT GCG GTT AAA GCA CCT GGC TTC GGC GAT CGT CGT AAA GCA ATG CTG CAG GAT ATC GCT ACC CTG ACC  902
Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Gln Asp Ile Ala Thr Leu Thr  296

GGT GGT ACC GTT ATC TCT GAA GAG ATC GGT ATG GAG CTG GAA AAA GCA ACT CTG GAA GAT CTG GGC CAG GCG AAA  977
Gly Gly Thr Val Ile Ser Glu Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly Gln Ala Lys  321

CGC GTT GTT ATC AAC AAA GAT ACC ACC ACC ATC ATC GAT GGC GTG GGC GAC GAA GCT GCA ATC CAG GGT CGC GTG 1052
Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp Gly Val Gly Asp Glu Ala Ala Ile Gln Gly Arg Val  346
```

Fig. 1A

```
ACT CAG ATT CGT CAG CAG ATC GAA GAA GCA ACT TCC GAC TAT GAC CGT GAA AAA CTG CAG GAG CGC GTA GCG AAA  1127
Thr Gln Ile Arg Gln Gln Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Val Ala Lys   371

CTG GCA GGC GGC GTT GCG GTT ATC AAA GTT GGT GCT GCG ACT GAA GTT GAA ATG AAA GAG AAG AAA GCC CGC GTT  1202
Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val   396

GAA GAT GCC CTG CAC GCT ACC CGT GCT GCG GTA GAA GAA GGC GTG GTT GCT GGT GGT GGC GTT GCG CTG ATT CGC  1277
Glu Asp Ala Leu His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly Val Ala Leu Ile Arg   421

GTA GCG TCT AAA ATT GCC GGC CTG AAA GGT CAG AAC GAA GAC CAG AAC GTA GGT ATC AAA GTT GCG CTG CGC GCA  1352
Val Ala Ser Lys Ile Ala Gly Leu Lys Gly Gln Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu Arg Ala   446

ATG GAA TCC CCA CTG CGT CAA ATC GTA CTG AAC TGC GGC GAA GAG CCG TCT GTA GTG GCT AAC ACC GTG AAA GCC  1427
Met Glu Ser Pro Leu Arg Gln Ile Val Leu Asn Cys Gly Glu Glu Pro Ser Val Val Ala Asn Thr Val Lys Ala   471

GGT GAC GGT AAC TAC GGT TAC AAC GCT GCA ACT GAA GAA TAC GGC AAC ATG ATC GAC ATG GGT ATC CTG GAT CCA  1502
Gly Asp Gly Asn Tyr Gly Tyr Asn Ala Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Met Gly Ile Leu Asp Pro   496

ACC AAA GTA ACT CGT TCT GCT CTG CAG TAC GCG GCT TCT GTT GCG GGT CTG ATG ATC ACC ACC GAG TGC ATG GTT  1577
Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala Ser Val Ala Gly Leu Met Ile Thr Thr Glu Cys Met Val   521

ACC GAC CTG CCG AAA GGC GAT GCA CCT GAC TTA GGT GCT GCT GGT GGT ATG GGC GGC ATG GGC GGA ATG ATG TGA  1652
Thr Asp Leu Pro Lys Gly Asp Ala Pro Asp Leu Gly Ala Ala Gly Gly Met Gly Gly Met Gly Gly Met Met *    546

TCAAGCC GAATTC                                                                                      1663
```

Fig. 1B

S. pneumoniae Hsp60-2 gene (SEQ ID NOS:3 and 4)

```
GAATTCGGCT TCAT ATG GCA AAA GAA ATT AAA TTT TCA TCA GAT GCC CGT TCA GCT ATG GTC CGT GGT GTC GAT ATC   77
                Met Ala Lys Glu Ile Lys Phe Ser Ser Asp Ala Arg Ser Ala Met Val Arg Gly Val Asp Ile   21

CTT GCA GAT ACT GTT AAA GTA ACT TTG GGA CCA AAA GGT CGC AAT GTC GTT CTT GAA AAG TCA TTC GGT TCA CCC  152
Leu Ala Asp Thr Val Lys Val Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Glu Lys Ser Phe Gly Ser Pro   46

TTG ATT ACC AAT GAC GGT GTG ACT ATT GCC AAA GAA ATT GAA TTA GAA GAC CAT TTT GAA AAT ATG GGT GCC AAA  227
Leu Ile Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp His Phe Glu Asn Met Gly Ala Lys   71

TTG GTA TCA GAA GTA GCT TCA AAA ACC AAT GAT ATC GCA GGT GAT GGA ACT ACA ACT GCA ACT GTT TTG ACC CAA  302
Leu Val Ser Glu Val Ala Ser Lys Thr Asn Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Thr Gln   96

GCA ATC GTC CGT GAA GGA ATC AAA AAC GTC ACA GCA GGT GCA AAT CCA ATC GGT ATT CGT CGT GGG ATT GAA ACA  377
Ala Ile Val Arg Glu Gly Ile Lys Asn Val Thr Ala Gly Ala Asn Pro Ile Gly Ile Arg Arg Gly Ile Glu Thr  121

GCA GTT GCC GCA GCA GTT GAA GCT TTG AAA AAC AAC GTC ATC CCT GTT GCC AAT AAA GAA GCT ATC GCT CAA GTT  452
Ala Val Ala Ala Ala Val Glu Ala Leu Lys Asn Asn Val Ile Pro Val Ala Asn Lys Glu Ala Ile Ala Gln Val  146

GCA GCC GTA TCT TCT CGT TCT GAA AAA GTT GGT GAG TAC ATC TCT GAA GCA ATG GAA AAA GTT GGC AAA GAC GGT  527
Ala Ala Val Ser Ser Arg Ser Glu Lys Val Gly Glu Tyr Ile Ser Glu Ala Met Glu Lys Val Gly Lys Asp Gly  171

GTC ATC ACC ATC GAA GAG TCA CGT GGT ATG GAA ACA GAG CTT GAA GTC GTA GAA GGA ATG CAG TTT GAC CGT GGT  602
Val Ile Thr Ile Glu Glu Ser Arg Gly Met Glu Thr Glu Leu Glu Val Val Glu Gly Met Gln Phe Asp Arg Gly  196

TAC CTT TCA CAG TAC ATG GTG ACA GAT AGC GAA AAA ATG GTG GCT GAC CTT GAA AAT CCG TAC ATT TTG ATT ACA  677
Tyr Leu Ser Gln Tyr Met Val Thr Asp Ser Glu Lys Met Val Ala Asp Leu Glu Asn Pro Tyr Ile Leu Ile Thr  221

GAC AAG AAA ATT TCC AAT ATC CAA GAA ATC TTG CCA CTT TTG GAA AGC ATT CTC CAA AGC AAT CGT CCA CTC TTG  752
Asp Lys Lys Ile Ser Asn Ile Gln Glu Ile Leu Pro Leu Leu Glu Ser Ile Leu Gln Ser Asn Arg Pro Leu Leu  246

ATT ATT GCG GAT GAT GTG GAT GGT GAG GCT CTT CCA ACT CTT GTT TTG AAC AAG ATT CGT GGA ACC TTC AAC GTA  827
Ile Ile Ala Asp Asp Val Asp Gly Glu Ala Leu Pro Thr Leu Val Leu Asn Lys Ile Arg Gly Thr Phe Asn Val  271

GTA GCA GTC AAG GCA CCT GGT TTT GGT GAC CGT CGC AAA GCC ATG CTT GAA GAT ATC GCC ATC TTA ACA GGC GGA  902
Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Glu Asp Ile Ala Ile Leu Thr Gly Gly  296

ACA GTT ATC ACA GAA GAC CTT GGT CTT GAG TTG AAA GAT GCG ACA ATT GAA GCT CTT GGT CAA GCA GCG AGA GTG  977
Thr Val Ile Thr Glu Asp Leu Gly Leu Glu Leu Lys Asp Ala Thr Ile Glu Ala Leu Gly Gln Ala Ala Arg Val  321

ACC GTG GAC AAA GAT AGC ACG GTT ATT GTA GAA GGT GCA GGA AAT CCT GAA GCG ATT TCT CAC CGT GTT GCG GTT 1052
Thr Val Asp Lys Asp Ser Thr Val Ile Val Glu Gly Ala Gly Asn Pro Glu Ala Ile Ser His Arg Val Ala Val  346
```

Fig. 2A

```
ATC AAG TCT CAA ATC GAA ACT ACA ACT TCT GAA TTT GAC CGT GAA AAA TTG CAA GAA CGC TTG GCC AAA TTG TCA  1127
Ile Lys Ser Gln Ile Glu Thr Thr Thr Ser Glu Phe Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala Lys Leu Ser   371

GGT GGT GTA GCG GTT ATT AAG GTC GGA GCC GCA ACT GAA ACT GAG TTG AAA GAA ATG AAA CTC CGC ATT GAA GAT  1202
Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu Thr Glu Leu Lys Glu Met Lys Leu Arg Ile Glu Asp   396

GCC CTC AAC GCT ACT CGT GCA GCT GTT GAA GAA GGT ATT GTT GCA GGT GGT GGA ACA GCT CTT GCC AAT GTG ATT  1277
Ala Leu Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Thr Ala Leu Ala Asn Val Ile   421

CCA GCT GTT GCT ACC TTG GAA TTG ACA GGA GAT GAA GCA ACA GGA CGT AAT ATT GTT CTC CGT GCT TTG GAA GAA  1352
Pro Ala Val Ala Thr Leu Glu Leu Thr Gly Asp Glu Ala Thr Gly Arg Asn Ile Val Leu Arg Ala Leu Glu Glu   446

CCT GTT CGT CAA ATT GCT CAC AAT GCA GGA TTT GAA GGA TCT ATC GTT ATC GAT CGT TTG AAA AAT GCT GAG CTT  1427
Pro Val Arg Gln Ile Ala His Asn Ala Gly Phe Glu Gly Ser Ile Val Ile Asp Arg Leu Lys Asn Ala Glu Leu   471

GGT ATA GGA TTC AAC GCA GCA ACT GGC GAG TGG GTT AAC ATG ATT GAT CAA GGT ATC ATT GAT CCA GTT AAA GTG  1502
Gly Ile Gly Phe Asn Ala Ala Thr Gly Glu Trp Val Asn Met Ile Asp Gln Gly Ile Ile Asp Pro Val Lys Val   496

AGT CGT TCA GCC CTA CAA AAT GCA GCA TCT GTA GCC AGC TTG ATT TTG ACA ACA GAA GCA GTC GTA GCC AAT AAA  1577
Ser Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu Ile Leu Thr Thr Glu Ala Val Val Ala Asn Lys   521

CCA GAA CCA GTA GCC CCA GCT CCA GCA ATG GAT CCA AGT ATG ATG GGT GGA ATG GGC GGA TGA TCAAAGC CGAATTC  1654
Pro Glu Pro Val Ala Pro Ala Pro Ala Met Asp Pro Ser Met Met Gly Gly Met Gly Gly *                     542
```

Fig. 2B

S. pyogenes Hsp60-1 gene (SEQ ID NOS: 5 and 6)

```
GAATTCGGCT TCAT ATG GCG GCT AAA GAT GTA AAA TTC GGT AAC GAC GCT CGT GTA AAA ATG CTC CGC GGC GTA AAC   77
                Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met Leu Arg Gly Val Asn   21

GTA CTG GCA GAC GCA GTT AAA GTA ACC CTG GGC CCG AAA GGC CGT AAC GTA GTG CTG GAC AAA TCC TTC GGC GCG  152
Val Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala   46

CCA ACC ATC ACG AAA GAT GGT GTT TCT GTA GCA CGT GAA ATC GAG CTG GAA GAC AAG TTC GAA AAC ATG GGC GCG  227
Pro Thr Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile Glu Leu Glu Asp Lys Phe Glu Asn Met Gly Ala   71

CAG ATG GTG AAA GAA GTG GCC TCT AAA GCG AAC GAC GCT GCA GGC GAC GGT ACC ACC ACC GCG ACC GTG CTG GCT  302
Gln Met Val Lys Glu Val Ala Ser Lys Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala   96

CAG GCT ATC ATC ACC GAA GGT CTG AAA GCC GTT GCT GCG GGC ATG AAC CCA ATG GAT CTG AAA CGT GGT ATC GAC  377
Gln Ala Ile Ile Thr Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn Pro Met Asp Leu Lys Arg Gly Ile Asp  121

AAA GCT GTC GCG TCC GCT GTT GAA GAA CTG AAA GCG CTG TCC GTA CCG TGC TCT GAC TCT AAA GCC ATT GCT CAG  452
Lys Ala Val Ala Ser Ala Val Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp Ser Lys Ala Ile Ala Gln  146

GTA GGT ACC ATC TCC GCT AAC TCC GAC GAA ACC GTA GGT AAA CTG ATC GCG GAA GCG ATG GAT AAA GTC GGT AAA  527
Val Gly Thr Ile Ser Ala Asn Ser Asp Glu Thr Val Gly Lys Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys  171

GAA GGC GTG ATC ACC GTT GAA GAC GGT ACC GGT CTG GAA GAC GAA CTG GAC GTG GTT GAA GGT ATG CAG TTC GAC  602
Glu Gly Val Ile Thr Val Glu Asp Gly Thr Gly Leu Glu Asp Glu Leu Asp Val Val Glu Gly Met Gln Phe Asp  196

CGC GGT TAC CTG TCC CCA TAC TTC ATC AAC AAG CCA GAA ACT GGC GCT GTT GAG CTG GAA AGC CCG TTC ATC CTG  677
Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Lys Pro Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu  221

CTG GCT GAC AAG AAA ATC TCC AAC ATC CGC GAA ATG CTG CCA GTG CTG GAA GCC GTT GCG AAA GCA GGC AAA CCG  752
Leu Ala Asp Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu Glu Ala Val Ala Lys Ala Gly Lys Pro  246

CTG GTT ATC ATT GCT GAA GAC GTT GAA GGC GAA GCG CTG GCG ACC CTG GTG GTT AAC ACC ATG CGT GGC ATC GTG  827
Leu Val Ile Ile Ala Glu Asp Val Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile Val  271

AAA GTG GCT GCG GTT AAA GCA CCT GGC TTC GGC GAC CGC CGT AAA GCG ATG CTG CAG GAT ATC GCT ACC CTG ACC  902
Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Gln Asp Ile Ala Thr Leu Thr  296

GGC GGT ACC GTC ATC TCT GAA GAG ATC GGT ATG GAG CTG GAA AAA GCG ACC CTG GAA GAC CTG GGC CAG GCT AAA  977
Gly Gly Thr Val Ile Ser Glu Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly Gln Ala Lys  321

CGT GTT GTG ATC AAC AAA GAC ACC ACC ACC ATC ATC GAT GGC GTG GGC GAC GAA GCG GCG ATT CAG GGC CGT GTT 1052
Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp Gly Val Gly Asp Glu Ala Ala Ile Gln Gly Arg Val  346
```

Fig. 3A

```
GGT CAG ATC CGT AAG CAG ATC GAA GAA GCC ACT TCC GAT TAC GAC CGT GAA AAA CTG CAG GAG CGC GTA GCG AAA 1127
Gly Gln Ile Arg Lys Gln Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Val Ala Lys  371

CTG GCA GGC GGT GTT GCG GTA ATC AAA GTC GGT GCT GCG ACT GAA GTT GAA ATG AAA GAG AAA AAA GCA CGC GTT 1202
Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val  396

GAC GAT GCC CTG CAC GCG ACC CGT GCT GCG GTA GAA GAA GGC GTG GTT GCT GGT GGT GGT GTG GCG CTG GTG CGT 1277
Asp Asp Ala Leu His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly Val Ala Leu Val Arg  421

GTT GCC GCG AAA CTG TCC GGC CTG ACT GCT CAG AAC GAA GAT CAG AAC GTG GGT ATC AAA GTT GCG CTG CGC GCA 1352
Val Ala Ala Lys Leu Ser Gly Leu Thr Ala Gln Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu Arg Ala  446

ATG GAA GCT CCA CTG CGT CAG ATC GTG TCC AAC GCC GGT GAA GAG CCA TCT GTT GTG ACC AAC AAC GTG AAA GCA 1427
Met Glu Ala Pro Leu Arg Gln Ile Val Ser Asn Ala Gly Glu Glu Pro Ser Val Val Thr Asn Asn Val Lys Ala  471

GGC GAA GGT AAC TAC GGT TAC AAC GCA GCA ACT GAA GAA TAC GGC AAC ATG ATC GAC TTC GGT ATC CTG GAT CCA 1502
Gly Glu Gly Asn Tyr Gly Tyr Asn Ala Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Phe Gly Ile Leu Asp Pro  496

ACC AAA GTG ACC CGT TCT GCT CTG CAG TAC GCG GCA TCT GTC GCT GGC CTG ATG ATC ACC ACC GAG TGC ATG GTG 1577
Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala Ser Val Ala Gly Leu Met Ile Thr Thr Glu Cys Met Val  521

ACC GAC CTG CCT AAA GGC GAC GCA CCT GAC TTA GGT GCT GCA GGC ATG GGT GGG ATG GGC GGT ATG ATG TGA TCAA 1653
Thr Asp Leu Pro Lys Gly Asp Ala Pro Asp Leu Gly Ala Ala Gly Met Gly Gly Met Gly Gly Met Met '       545

GCC GAATTC                                                                                      1662
```

Fig. 3B

S. pyogenes Hsp60-2 gene (SEQ ID NOS: 7 and 8)

```
GAATTCGGCT TCAT ATG GCA AAA GAA ATC AAA TTT TCA GCA GAT GCG CGT GCT GCC ATG GTG CGC GGA GTT GAT ATG   77
                Met Ala Lys Glu Ile Lys Phe Ser Ala Asp Ala Arg Ala Ala Met Val Arg Gly Val Asp Met   21

TTA GCA GAT ACC GTC AAA GTA ACG CTT GGT CCT AAA GGG CGC AAT GTT GTT CTT GAA AAA GCT TTT GGT TCT CCC  152
Leu Ala Asp Thr Val Lys Val Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Glu Lys Ala Phe Gly Ser Pro   46

TTA ATT ACT AAT GAC GGG GTA ACC ATT GCT AAA GAG ATC GAA TTA GAA GAT CAT TTT GAA AAC ATG GGA GCA AAA  227
Leu Ile Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp His Phe Glu Asn Met Gly Ala Lys   71

TTG GTG TCT GAA GTG GCT TCT AAA ACC AAT GAT ATT GCT GGT GAT GGG ACG ACT ACT GCA ACA GTT TTG ACA CAA  302
Leu Val Ser Glu Val Ala Ser Lys Thr Asn Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Thr Gln   96

GCC ATT GTT CAT GAA GGA CTA AAA AAT GTG ACA GCA GGT GCT AAT CCA ATT GGT ATC CGT CGA GGC ATT GAA ACA  377
Ala Ile Val His Glu Gly Leu Lys Asn Val Thr Ala Gly Ala Asn Pro Ile Gly Ile Arg Arg Gly Ile Glu Thr  121

GCA ACA GCA ACA GCT GTT GAA GCC TTG AAA GCC ATT GCT CAA CCT GTA TCT GGC AAG GAA GCT ATT GCT CAG GTC  452
Ala Thr Ala Thr Ala Val Glu Ala Leu Lys Ala Ile Ala Gln Pro Val Ser Gly Lys Glu Ala Ile Ala Gln Val  146

GCT GCA GTA TCA TCA CGC TCT GAA AAA GTT GGA GAG TAT ATC TCA GAA GCT ATG GAG CGT GTG GGC AAC GAT GGT  527
Ala Ala Val Ser Ser Arg Ser Glu Lys Val Gly Glu Tyr Ile Ser Glu Ala Met Glu Arg Val Gly Asn Asp Gly  171

GTG ATT ACC ATC GAA GAA TCT CGA GGT ATG GAA ACA GAA CTT GAA GTG GTT GAA GGC ATG CAA TTT GAC CGT GGT  602
Val Ile Thr Ile Glu Glu Ser Arg Gly Met Glu Thr Glu Leu Glu Val Val Glu Gly Met Gln Phe Asp Arg Gly  196

TAC CTG TCT CAA TAC ATG GTC ACA GAC AAT GAA AAA ATG GTT GCA GAC CTT GAA AAC CCA TTT ATC TTA ATC ACG  677
Tyr Leu Ser Gln Tyr Met Val Thr Asp Asn Glu Lys Met Val Ala Asp Leu Glu Asn Pro Phe Ile Leu Ile Thr  221

GAT AAA AAA GTG TCA AAC ATC CAA GAC ATT TTG CCA CTA CTT GAG GAA GTT CTT AAA ACC AAC CGT CCA TTA CTC  752
Asp Lys Lys Val Ser Asn Ile Gln Asp Ile Leu Pro Leu Leu Glu Glu Val Leu Lys Thr Asn Arg Pro Leu Leu  246

ATT ATT GCA GAT GAT GTG GAT GGT GAA GCA CTT CCA ACC CTT GTC TTG AAC AAG ATT CGT GGT ACT TTC AAT GTG  827
Ile Ile Ala Asp Asp Val Asp Gly Glu Ala Leu Pro Thr Leu Val Leu Asn Lys Ile Arg Gly Thr Phe Asn Val  271

GTT GCT GTC AAA GCG CCA GGA TTT GGT GAT CGT CGT AAA GCT ATG CTT GAA GAC ATT GCT ATC TTG ACA GGT GGT  902
Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Glu Asp Ile Ala Ile Leu Thr Gly Gly  296

ACA GTG ATT ACA GAG GAT CTA GGA CTT GAA TTA AAA GAT GCT ACA ATG ACA GCC CTT GGA CAG GCT GCT AAG ATT  977
Thr Val Ile Thr Glu Asp Leu Gly Leu Glu Leu Lys Asp Ala Thr Met Thr Ala Leu Gly Gln Ala Ala Lys Ile  321

ACA GTT GAT AAA GAT AGC ACA GTA ATT GTT GAA GGT TCA GGA AGT TCA GAA GCT ATT GCT AAC CGT ATT GCA CTG 1052
Thr Val Asp Lys Asp Ser Thr Val Ile Val Glu Gly Ser Gly Ser Ser Glu Ala Ile Ala Asn Arg Ile Ala Leu  346
```

Fig. 4A

```
ATT AAA TCG CAA TTA GAA ACA ACA ACT TCT GAC TTT GAC CGT GAA AAA CTA CAA GAA CGT TTG GCG AAA TTA GCT  1127
Ile Lys Ser Gln Leu Glu Thr Thr Thr Ser Asp Phe Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala Lys Leu Ala   371

GGT GGT GTA GCT GTT ATC AAA GTA GGA GCT CCA ACA GAG ACA GCT TTA AAA GAA ATG AAA CTT CGC ATT GAG GAT  1202
Gly Gly Val Ala Val Ile Lys Val Gly Ala Pro Thr Glu Thr Ala Leu Lys Glu Met Lys Leu Arg Ile Glu Asp   396

GCT CTA AAT GCT ACA CGT GCA GCC GTT GAA GAA GGT ATC GTT GCT GGT GGT GGA ACA GCA CTT ATT ACG GTT ATT  1277
Ala Leu Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Thr Ala Leu Ile Thr Val Ile   421

GAA AAA GTA GCA GCT CTT GAG CTT GAG GGC GAT GAT GCT ACT GGA CGT AAC ATT GTG CTT CGT GCT CTA GAA GAG  1352
Glu Lys Val Ala Ala Leu Glu Leu Glu Gly Asp Asp Ala Thr Gly Arg Asn Ile Val Leu Arg Ala Leu Glu Glu   446

CCT GTA CGT CAA ATT GCT TTA AAT GCT GGG TAC GAA GGC TCC GTA GTT ATT GAC AAG TTG AAA AAC AGC CCT GCA  1427
Pro Val Arg Gln Ile Ala Leu Asn Ala Gly Tyr Glu Gly Ser Val Val Ile Asp Lys Leu Lys Asn Ser Pro Ala   471

GGA ACA GGA TTT AAT GCT GCA ACA GGT GAG TGG GTT GAT ATG ATT AAA ACA GGA ATC ATT GAC CCT GTC AAA GTA  1502
Gly Thr Gly Phe Asn Ala Ala Thr Gly Glu Trp Val Asp Met Ile Lys Thr Gly Ile Ile Asp Pro Val Lys Val   496

ACA CGA TCA GCG CTT CAA AAT GCA GCT TCT GTA GCT AGT CTT ATT TTG ACA ACA GAA GCA GTT GTT GCT AAT AAA  1577
Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu Ile Leu Thr Thr Glu Ala Val Val Ala Asn Lys   521

CCT GAA CCA GCT ACG CCA GCG CCA GCA ATG CCA GCA GGT ATG GAT CCA GGA ATG ATG GGT GGG ATG GGC GGA TAA  1652
Pro Glu Pro Ala Thr Pro Ala Pro Ala Met Pro Ala Gly Met Asp Pro Gly Met Met Gly Gly Met Gly Gly *     546

GCCGAAT TC                                                                                          1661
```

Fig. 4B

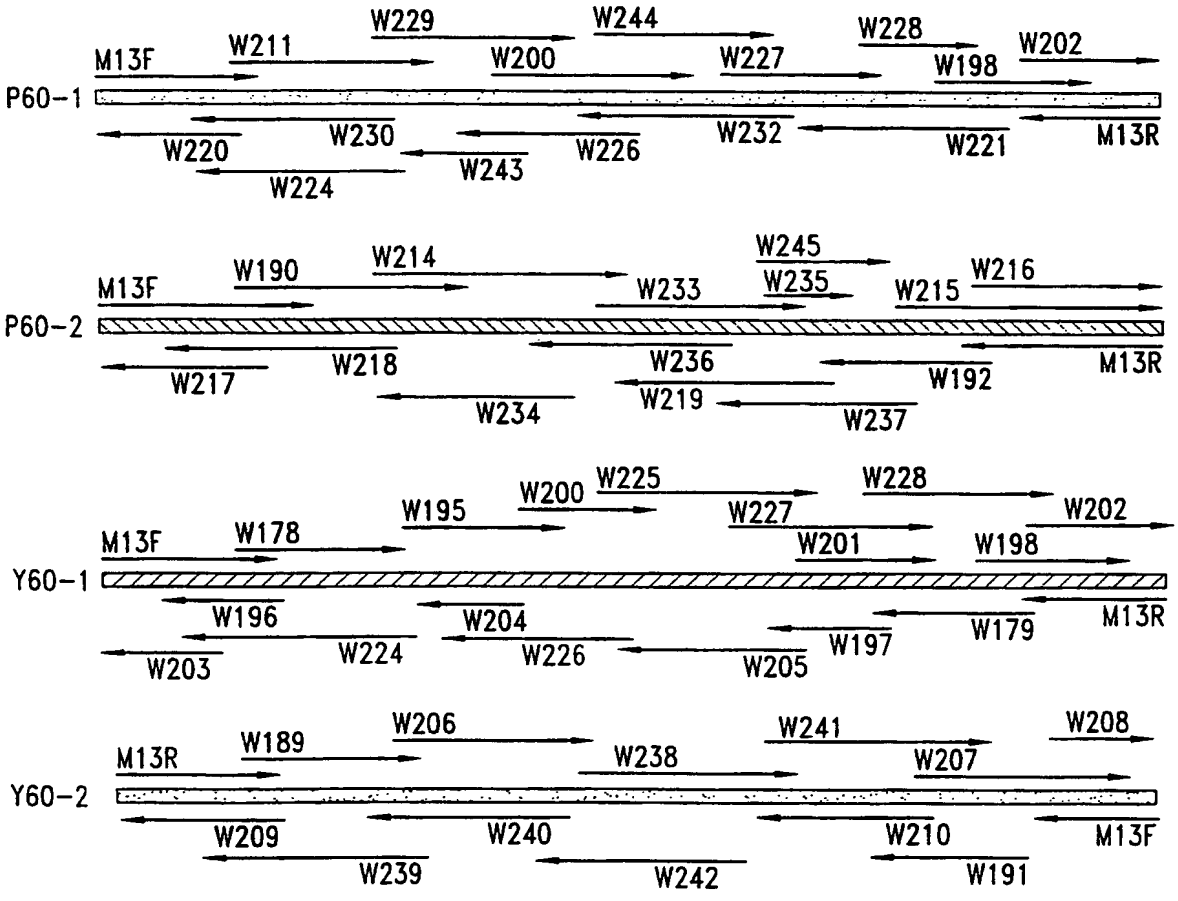

Sequencing strategy (scale: 1cm=approx. 100bp)
P60-1
M13F
W211
W229
W244
W200
W227
W228
W198
W202
W220
W230
W243
W226
W232
W221
M13R
W224
P60-2
M13F
W190
W214
W233
W245
W235
W215
W216
W217
W218
W234
W236
W219
W192
M13R
W237
Y60-1
M13F
W178
W195
W200
W225
W227
W228
W201
W198
W202
W196
W204
W203
W224
W226
W205
W197
W179
M13R
Y60-2
M13R
W189
W206
W238
W241
W207
W208
W209
W240
W239
W242
W210
W191
M13F

Fig. 5

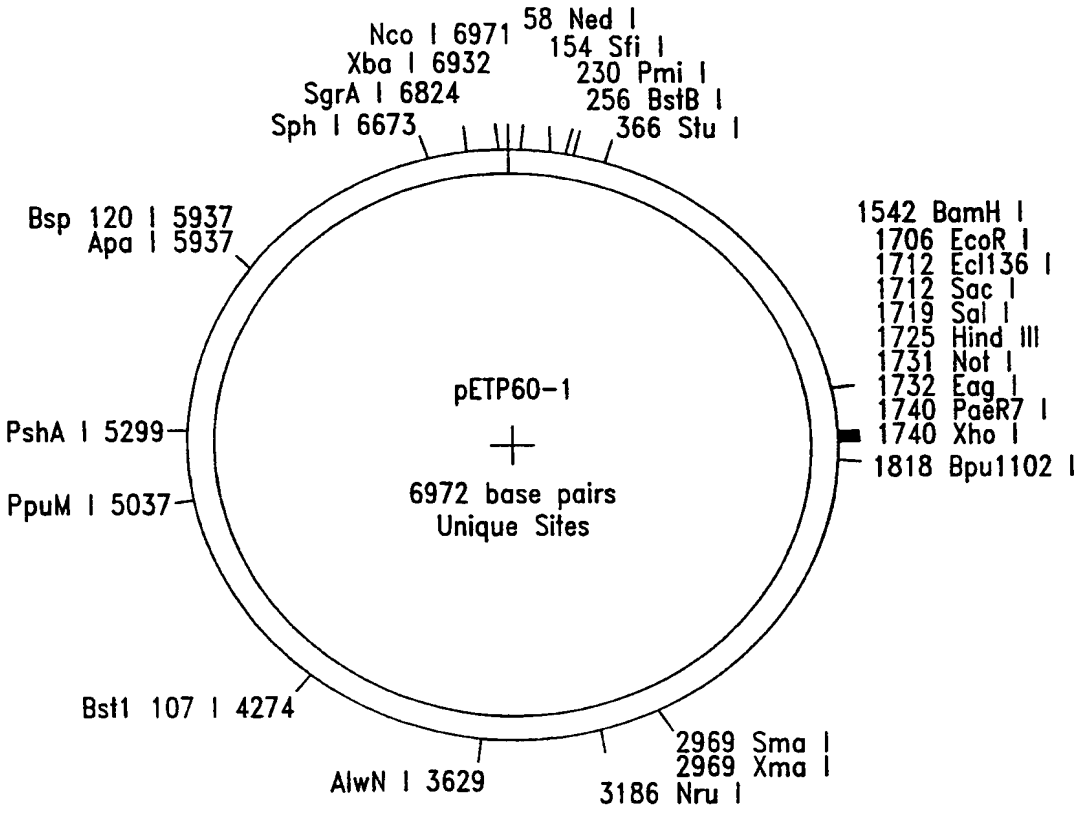

58 Ned I
Nco I 6971
154 Sfi I
Xba I 6932
230 Pmi I
SgrA I 6824
256 BstB I
Sph I 6673
366 Stu I
Bsp 120 I 5937
Apa I 5937
1542 BamH I
1706 EcoR I
1712 Ecl136 I
1712 Sac I
1719 Sal I
1725 Hind III
1731 Not I
1732 Eag I
1740 PaeR7 I
1740 Xho I
1818 Bpu1102 I
pETP60-1
6972 base pairs
Unique Sites
PshA I 5299
PpuM I 5037
Bst1 107 I 4274
AlwN I 3629
2969 Sma I
2969 Xma I
3186 Nru I

Fig. 6

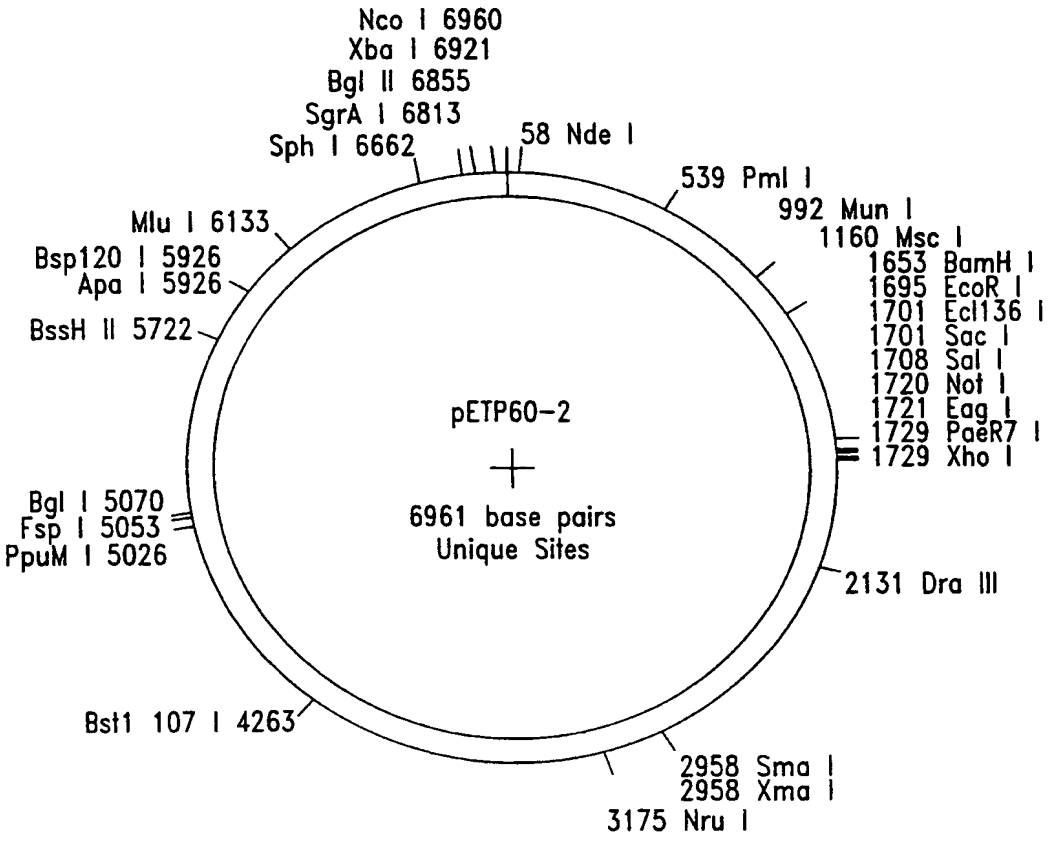

Nco I 6960
Xba I 6921
Bgl II 6855
SgrA I 6813
Sph I 6662
58 Nde I
539 PmI I
992 Mun I
1160 Msc I
1653 BamH I
1695 EcoR I
1701 Ecl136 I
1701 Sac I
1708 Sal I
1720 Not I
1721 Eag I
1729 PaeR7 I
1729 Xho I
2131 Dra III
2958 Sma I
2958 Xma I
3175 Nru I
Bst1 107 I 4263
Bgl I 5070
Fsp I 5053
PpuM I 5026
BssH II 5722
Bsp120 I 5926
Apa I 5926
Mlu I 6133
pETP60-2
6961 base pairs
Unique Sites

Fig. 7

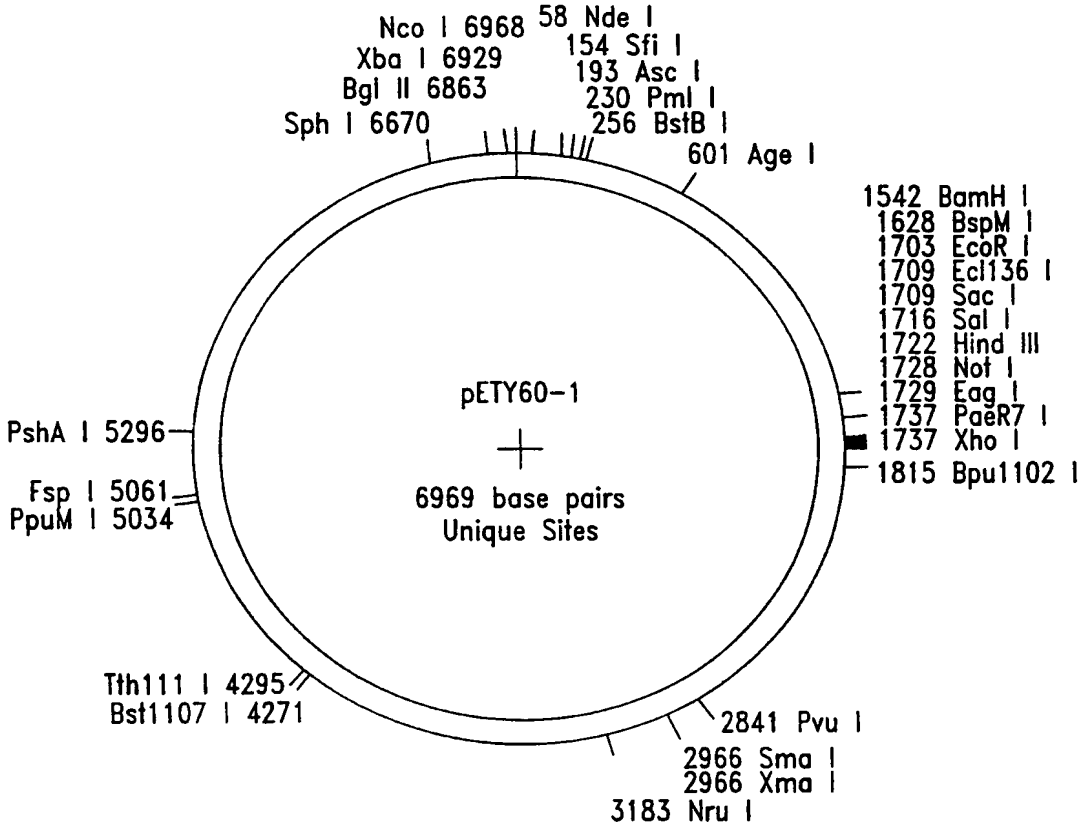

Nco I 6968
58 Nde I
154 Sfi I
Xba I 6929
193 Asc I
Bgl II 6863
230 PmI I
Sph I 6670
256 BstB I
601 Age I
1542 BamH I
1628 BspM I
1703 EcoR I
1709 Ecl136 I
1709 Sac I
1716 Sal I
1722 Hind III
1728 Not I
1729 Eag I
1737 PaeR7 I
1737 Xho I
1815 Bpu1102 I
pETY60-1
6969 base pairs
Unique Sites
PshA I 5296
Fsp I 5061
PpuM I 5034
Tth111 I 4295
Bst1107 I 4271
2841 Pvu I
2966 Sma I
2966 Xma I
3183 Nru I

Fig. 8

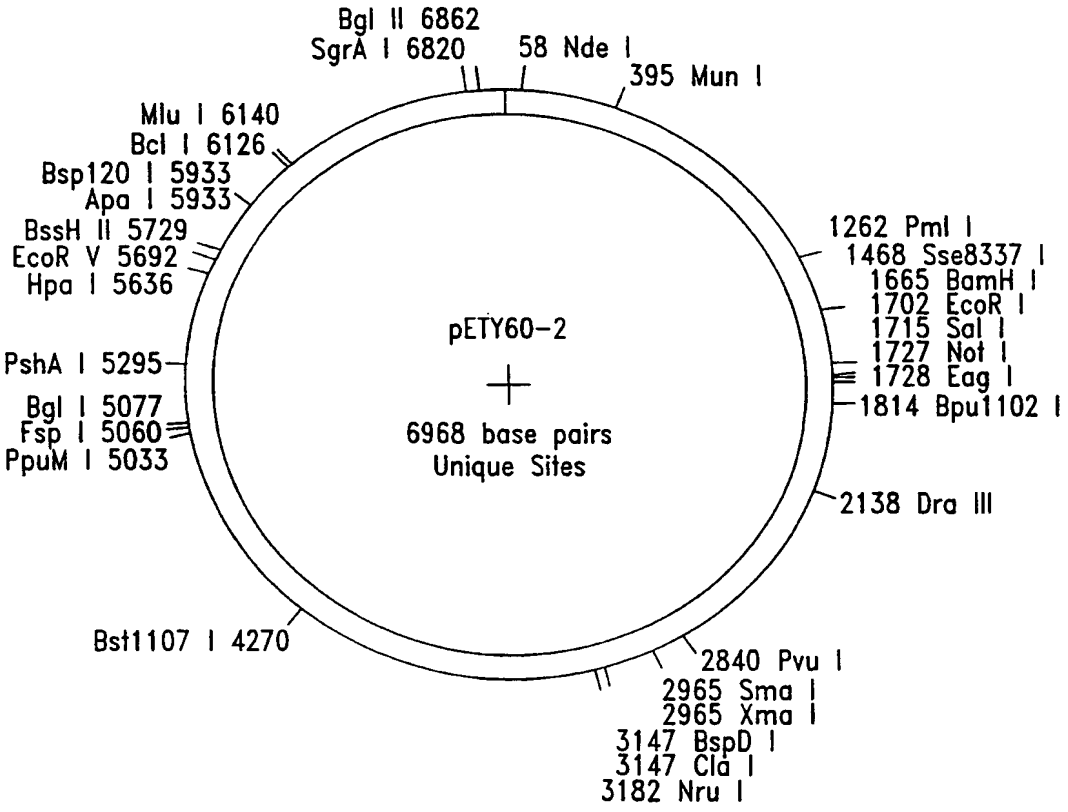

Bgl II 6862
SgrA I 6820
58 Nde I
395 Mun I
Mlu I 6140
Bcl I 6126
Bsp120 I 5933
Apa I 5933
BssH II 5729
EcoR V 5692
Hpa I 5636
PshA I 5295
Bgl I 5077
Fsp I 5060
PpuM I 5033
Bst1107 I 4270
1262 Pml I
1468 Sse8337 I
1665 BamH I
1702 EcoR I
1715 Sal I
1727 Not I
1728 Eag I
1814 Bpu1102 I
2138 Dra III
2840 Pvu I
2965 Sma I
2965 Xma I
3147 BspD I
3147 Cla I
3182 Nru I
pETY60-2
6968 base pairs
Unique Sites

Fig. 9

```
MA-----------------------------------------------KDIKFGEEARRAMLRGVNALADAVKVTLGPKGRNVVLEKSFGAPTITKDGVTVAKEIELEDKFENMGAQLVKEVASKTNDV
            10        20        30        40        50        60        70        80        90       100       110       120       130

S.pneumoniae  MA----------------------------------------------AKDVKFGNDARVKMLRGVNVLADAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAREIELEDKFENMGAQMVKEVASKANDA 84
S.pyogenes hs MA----------------------------------------------AKDVKFGNDARVKMLRGVNVLADAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAREIELEDKFENMGAQMVKEVASKANDA 84
S.pneumoniae  MA-----------------------------------------------KEIKFSSDARSAMVRGVDILADTVKVTLGPKGRNVVLEKSFGSPLITNDGVTIAKEIELEDHFENMGAKLVSEVASKTNDI 84
S.pyogenes hs MA-----------------------------------------------KEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIELEDHFENMGAKLVSEVASKTNDI 83
B.subtilis gr MA-----------------------------------------------KEIKFSEEARRAMLRGVDALADAVKVTLGPKGRNVVLEKKFGSPLITNDGVTIAKEIELEDAFENMGAKLVAEVASKTNDV 83
Clostridium h MA-----------------------------------------------KTLLFGEEARRSMQAGVDKLANTVKVTLGPKGRNVILDKKFGSPLITNDGVTIAREIELEDAYENMGAQLVKEVATKTNDV 83
Cowdria hsp60 MANM-----------------------------------------------VVTGEQLDKSIREVVRILEDAVGCTAGPKGLTVAISKPYGAPEVTKDGYKVMKSIKPEDPLALAIANIIAQSASQCNDK 83
Haemophilus h MA----------------------------------------------AKDVKFGNDARVKMLNGVNILADAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAREIELEDKFENMGAQMVKEVASKANDA 84
L.pneumophila MA-----------------------------------------------KELRFGDDARLQMLAGVNALADAVQVTMGPRGRNVVLEKSYGAPTVTKDGVSVAKEIEFEHRFMNMGAQMVKEVASKTSDT 83
M.avium hsp60 MA-----------------------------------------------KTIAYDEEARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEIELEDPYEKIGAELVKEVAKKTDDV 83
M.bovis hsp60 MA-----------------------------------------------KTIAYDEEARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEIELEDPYEKIGAELVKEVAKKTDDV 83
M.leprae groE MS-----------------------------------------------KLIEYDETARHAMEVGMNKLADTVRVTLGPRGRHVVLAKAFGGPTITNDGVTVAREIDLEDPFENLGAQLVKSVATKTNDV 83
M.leprae hsp6 MA-----------------------------------------------KTIAYDEEARRGLERGLNSLADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEIELEDPYEKIGAELVKEVAKKTDDV 83
M.tuberculosi MS-----------------------------------------------KLIEYDETARRAMEVGMDKLADTVRVTLGPRGRHVVLAKAFGGPTVTNDGVTVAREIELEDPFEDLGAQLVKSVATKTNDV 83
N.meningitidi MA----------------------------------------------AKDVQFGNEVRQKMVNGVNILANAVRVTLGPKGRNVVVDRAFGGPHITKDGVTVAKEIELKDKFENMGAQMVKEVASKTNDV 84
S.aureus hsp6 MV-----------------------------------------------KQLKFSEDARQAMLRGVDQLANAVKVTIGPKGRNVVLDKEFTAPLITNDGVTIAKEIELEDPYENMGAKLVQEVANKTNEI 83
Synechocystis MA-----------------------------------------------KSIIYNDEARRALERGMDILAEAVAVTLGPKGRNVVLEKKFGSPQIINDGITIAKEIELEDHVENTGVSLIRQAASKTNDV 83
Synechocystis MS-----------------------------------------------KLISFKDESRRSLEAGINALADAVRITLGPKGRNVLLEKQYGAPQIVNDGITVAKEIELSNPEENAGAKLIQEVASKTKEI 83
Tsukamurella  MA-----------------------------------------------KTIAFDKKARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEIELEDPYEKIGAELVKEVAKKTDDV 83
S.pombe hsp60 M-----VSFLSS--SVSRLPLRIAGRRIPGRFAVPQVRT---Y---A--KDLKFGVDARASLLTGVDTLARAVSVTLGPKGRNVLIDQPFGSPKITKDGVTVARSVSLKDKFENLGARLVQDVASKTNEV 115
S.cerevisiae  M-------------------LRSSVVRSRATLRPLLRRA---Y---SSHKELKFGVEGRASLLKGVETLAEAVAATLGPKGRNVLIEQPFGPPKITKDGVTVAKSIVLKDKFENMGAKLLQEVASKTNEA 105
P.falciparum  M-----IS---TLRGKI-FNNGSNRNKCVSILSNIQKRN---I---S--KDIRFGSDARTAMLTGCNKLADAVSVTLGPKGRNVIIEQSFGSPKITKDGVTVAKSIEFNNKLANLGAQMVKQVAANTNGK 113
Onchocerca hs MTNV-----------------------------------------------VVSGEQLQQAFREVAAVIDSTVAVTAGPRGKTVGINKPYGAPEITKDGYKVMKGIKPEKPLNAAITSIFAQSCSQCNDK 83
C.elegans hsp M-----LRLA---------RKG---------LQTAVVRS---Y---A--KDVKFGAEGRQAMLVGVNLLADAVSVTMGPKGRNVIIEQSWGSPKITKDGVTVAKSIDLKDKYQNLGAKLIQDVANKANEE 99
D.melanogaste M-----FRLPVSL-ARSSISRQ---------LA---MRG---Y---A--KDVRFGPEVRAMMLQGVDVLADAVAVTMGPKGRNVIIEQSVGLAKITKDGVTVAKSIELKDKFQNIGAKLVQDLANNTNEE 104
human hsp60   M-----LRLPTVFRQMRPVSRV---------LAPHLTRA---Y---A--KDVKFGADARALMLQGVDLLADAVAVTMGPKGRTVIIEQGWGSPKVTKDGVTVAKSIDLKDKYKNIGAKLVQDVANNTNEE 108
Arabidopsis h M-----YRFASNLASKARIAQN---ARQVSS-RMSWSRN---Y---AA-KEIKFGVEARALMLKGVEDLADAVKVTMGPKGRNVVIEQSWGAPKVTKDGVTVAKSIEFKDKIKNVGASLVKQVANATNDV 114
maize hsp60   M-----YRAAASLASKARQAGSSSAARQVGS-RLAWSRN---Y---AA-KDIKFGVEARALMLRGVEELADAVKVTMGPKGRNVVIEQSFGAPKVTKDGVTVAKSIEFKDRVKNVGASLVKQVANATNDN 117
RUBISCO chape MASTNALSSTSILRSPTNQAQTSLSKKVKQHGRVNFRQKPNRFVVKAAAKDIAFDQHSRSAMQAGIDKLADAVGLTLGPRGRNVVLD-EFGSPKVVNDGVTIARAIELPDPMENAGAALIREVASKTNDS 129
```

Fig. 10A

```
              AGDGTTTATVLAQAIVKEGLKNVAAGANPMDLRRGIDKAVDAVVEELKAIAKPV--ETKEEIAQVATISANGDEEIGE---LIAEAMEKVGKEGVITV---EEGKTLETELEVVEG-MQFDRGYISPYFI
                   140       150       160       170       180       190       200       210       220       230       240       250       260

S.pneumoniae  AGDGTTTATVLAQSIITEGLKAVAAGMNPMDLKRGIDKAVAAAVEELKALSVPC--SDSKAIAQVGTISANSDETVGK---LIAEAMDKVGKEGVITV---EDGTGLQDELDVVEG-MQFDRGYLSPYFI 205
S.pyogenes hs AGDGTTTATVLAQAIITEGLKAVAAGMNPMDLKRGIDKAVASAVEELKALSVPC--SDSKAIAQVGTISANSDETVGK---LIAEAMDKVGKEGVITV---EDGTGLEDELDVVEG-MQFDRGYLSPYFI 205
S.pneumoniae  AGDGTTTATVLTQAIVREGIKNVTAGANPIGIRRGIETAVAAAVEALKNNVIPV--ANKEAIAQVAAVSS-RSEKVGE---YISEAMEKVGKDGVITI---EESRGMETELEVVEG-MQFDRGYLSQYMV 203
S.pyogenes hs AGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIETATATAVEALKAIAQPV--SGKEAIAQVAAVSS-RSEKVGE---YISEAMERVGNDGVITI---EESRGMETELEVVEG-MQFDRGYLSQYMV 203
B.subtilis gr AGDGTTTATVLAQAMIREGLKNVTAGANPVGVRKGMEQAVAVAIENLKEISKPI--EGKESIAQVAAISA-ADEEVGS---LIAEAMERVGNDGVITI---EESKGFTTELEVVEG-MQFDRGYASPYMV 203
Clostridium h AGDGTTTATLLAQAIIREGLKNVTAGANPILIRNGIKTAVEKAVEEIQKISKPV--NGKEDIARVAAISA-ADEKIGK---LIADAMEKVGNEGVITV---EESKSMGTELDVVEG-MQFDRGYVSAYMV 203
Cowdria hsp60 VGDGTTTCSILTAKVIEEVSKVKAAGADIICVREGVLKAKEAVLEALKCMKREVLSE--EEIAQVATISANGDKNIGTK---IAQCVKEVGKDGVITVEESKGFKEL--DVEKTDG-MQFDRGYLSPYFV 205
Haemophilus h AGDGTTTATVLAQAIVNEGLKAVAAGMNPMDLKRGIDKAVNSVVAELKNLSKPC--ETSKEIEQVGTISANSDSIVGQ---LIAQAMEKVGKEGVITV---EDGTGLEDELDVVEG-MQFDRGYLSPYFI 205
L.pneumophila AGDGTTTATVLARSILVEGHKAVAAGMNPMDLKRGIDKAVLAVTKKLQAMSKPC--KDSKAIAQVGTISANSDEAIGA---IIAEAMEKVGKEGVITV---EDGNGLENELSVVEG-MQFDRGYISPYFI 204
M.avium hsp60 AGDGTTTATVLAQALVREGLRNVAAGANPLGLKRGIEKAVEKVTETLLKSAKEV--ETKDQIAATAAISA-GDQSIGD---LIAEAMDKVGNEGVITV---EESNTFGLQLELTEG-MRFDKGYISGYFV 203
M.bovis hsp60 AGDGTTTATVLAQALVREGLRNVAAGANPLGLKRGIEKAVEKVTETLLKGAKEV--ETKEQIAATAAISA-GDQSIGD---LIAEAMDKVGNEGVITV---EESNTFGLQLELTEG-MRFDKGYISGYFV 203
M.leprae groE AGDGTTTATVLAQALVKGGLRMVAAGANPVALGAGISKAADAVSEALLAVATPV--AGKDAITQVATVSSR-DEQIGA---LVGEGMNKVGTDGVVSV---EESSTLDTELEFTEG-VGFDKGFLSAYFV 203
M.leprae hsp6 AGDGTTTATVLAQALVKEGLRNVAAGANPLGLKRGIEKAVDKVTETLLKDAKEV--ETKEQIAATAAISA-GDQSIGD---LIAEAMDKVGMEGVITV---EESNTFGLQLELTEG-MRFDKGYISGYFV 203
M.tuberculosi AGDGTTTATILAQALIKGGLRLVAAGVNPIALGVGIGKAADAVSEALLASATPV--SGKTGIAQVATVSSR-DEQIGD---LVGEAMSKVGHDGVVSV---EESSTLGTELEFTEG-IGFDKGFLSAYFV 203
N.meningitidi AGDGTTTATVLAQSIVAEGMKYVTAGMNPTDLKRGIDKAVAALVEELKNIAKPC--DTSKEIAQVGSISANSDEQVGA---IIAEAMEKVGKEGVITV---EDGKSLENELDVVEG-MQFDRGYLSPYFI 205
S.aureus hsp6 AGDGTTTATVLAQAMIQEGLKNVTSGANPVGLRQGIDKAVKVAVEALHENSQKV--ENKNEIAQVGAISA-ADEEIGR---YISEATEKVGNDGVITIITIEESNRLNTELEL--G-MQFDRGYQSPYMV 204
Synechocystis AGDGTTTATVLAHAIVKEGLRNVAAGANPISLKRGIDKATDFLVARIKEHAQPV--GDSKAIAQVGAISAGNDEEVGQ---MIANAMDKVGQEGVISL---EEGKSMTTELEITEG-MRFDKGYISPYFV 204
Synechocystis AGDGTTTATIIAQALVREGLRNVAAGANPVALRRGIEKVTTFLVQEIEAVAKPV---EGSAIAQVATVSSGNDPEVGA---MIADAMDKVTKDGVITV---EESKSLNTELEVVEG-MQIDRGYISPYFI 203
Tsukamurella  AGDGTTTATVLAQALVREGLRNVAAGANPLGLKRGIEKAVEAVTEHLLKAAKEV--ETKDQIAATAGISA-GDPAIGE---LIAEAMDKVGKEGVITV---EESNTFGLQLELTEG-MRFDKGFISGYFA 203
S.pombe hsp60 AGDGTTTATVLTRAIFSETVRNVAAGCNPMDLRRGIQLAVDNVVEFLQANKRDI--TTSEEISQVATISANGDTHIGE---LLAKAMERVGKEGVITV---KEGRTISDELEVTEG-MKFDRGYISPYFI 236
S.cerevisiae  AGDGTTSATVLGRAIFTESVKNVAAGCNPMDLRRGSQVAVEKVIEFLSANKKEI--TTSEEIAQVATISANGDSHVGK---LLASAMEKVGKEGVITI---REGRTLEDELEVTEG-MRFDRGFISPYFI 226
P.falciparum  AGDGTTTATILARSIFQQGCKAVDSGMNPMDLLRGINKGVEKVLEYLNSIKKDV--TTTEEIFNVASIS-NGDKNIGQ---LIADTMKKVGKEGTITV---TEGKTLQHELEIVEG-IKFDRGYISPYFI 233
Onchocerca hs VGDGTTTCSILTSGMIVEASKSIAAGNDRISIKNGMQKAKDVVLKEVASMARTISLEKIDEVAQVAIISANGDRSIGSN---IADAVKKVGKEGVITVEESKGSKEL--EVELTTG-MQFDRGYLSPYFI 207
C.elegans hsp AGDGTTCATVLTRAIAKEGFERHSSRGNAVEIRRGVMNAVEVVVAELKKISKKV--TTPEEIAQVATISANGDTVVGN---LISDAMKKVGTTGVITV---KDGKTLNDQLELIEG-MKFDRGYISPYFI 220
D.melanogaste AGDGTTTATFLARAIAKEGFEKISKGGNPVEIRRGVMLAVETVKDNLKTMSRPV--STPEEIAQVATISANGDREIGNGKVSVSEAMKKVGRDGVITV---KDGKTLTDELEVIEGTMRFDRGYISPYFI 229
human hsp60   AGDGTTTATVLARSIAKEGFEKISKGANPVEIRRGVMLAVDAVIAELKKQSKPV--TTPEEIAQVATISANGDKEIGN---IISDAMKKVGRKGVITV---KDGKTLNDELEIIEG-MKFDRGYISPYFI 229
Arabidopsis h AGDGTTCATVLTRAIFAEGCKSVAAGMNAMDLRRGISMAVDAVVTNLKSKARMI--STSEEIAQVGTISANGEREIGE---LIAKAMEKVGKEGVITI---QDGKTLFNELEVVEG-MKLDRGYTSPYFI 235
maize hsp60   AGDGTTCATVLTKAIFTEGCKSVAAGMNAMDLRRGISMAVDAVVTNLKGMARMI--STSEEIAQVGTISANGEREIGE---LIAKAMEKVGKEGVITI---ADGNTLYNELEVVEG-MKLDRGYISPYFI 238
RUBISCO chape AGDGTTTASILAREIIKLGLLNVTSGANPVSIKKGIDKTVAALVEELEKLARPV--KGGDDIKAVATISAGNDELIGK---MIAEAIDKVGPDGVLSI---ESSNSFETTVEVEEG-MEIDRGYISPQFV 250
```

Fig. 10B

```
                 TDSEKQKAELEDPLILLTDKKISNIQDLLPVLEEVA--QAGKPLLIIAEDVEGEALATLVVNKLRGTLKVVAVKAPGFGDRRKAMLQDIAILTGGQVISEE-LGLSLEDATLED-LGQAKKVVVTKDDTT
                        270       280       290       300       310       320       330       340       350       360       370       380       390

S.pneumoniae     NKPETGAVELESPFILLADKKISNIREMLPVLEAVA--KAGKPLLIIAEDVEGEALATLVVNTMRGIVKVAAVKAPGFGDRRKAMLQDIATLTGGTVISEE-IGMELEKATLED-LGQAKRVVINKDTTT 331
S.pyogenes hs    NKPETGAVELESPFILLADKKISNIREMLPVLEAVA--KAGKPLVIIAEDVEGEALATLVVNTMRGIVKVAAVKAPGFGDRRKAMLQDIATLTGGTVISEE-IGMELEKATLED-LGQAKRVVINKDTTT 331
S.pneumoniae     TDSEKMVADLENPYILITDKKISNIQEILPLLESIL--QSNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVITED-LGLELKDATIEA-LGQAARVTVDKDSTV 329
S.pyogenes hs    TDNEKMVADLENPFILITDKKVSNIQDILPLLEEVL--KTNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVITED-LGLELKDATMTA-LGQAAKITVDKDSTV 329
B.subtilis gr   TDSDKMEAVLDNPYILITDKKITNIQEILPVLEQVV--QQGKPLLLIAEDVEGEALATLVVNKLRGTFNAVAVKAPGFGDRRKAMLEDIAVLTGGEVITED-LGLDLKSTQIAQ-LGRASKVVVTKENTT 329
Clostridium h    TDTEKMEAVLDNPLVLITDKKISNIQDLLPLLEQIV--QAGKKLLIIADDIEGEAMTTLVVNKLRGTFTCVGVKAPGFGDRRKEMLQDIATLTGGVVISDE-VGGDLKEATLDM-LGEAESVKVTKESTT 329
Cowdria hsp60    TNSEKMLVEFENPYILLTEKKLNIIQPLLPILENIA--RSGRPLLIIAEDVEGEALSTLVLNKLRGGLHVAAVKAPGFGDRRKDMLGDIAILTGAKHVINDELAIKMEDLTLCD-LGTAKNIRITKDTTT 332
Haemophilus h    NKPETATVELDNPFILLVDKKISNIRELLPVLEGVA--KAGKPLLIIAEDVEGEALATLVVNTMRGIVKVAAVKAPGFGDRRKAMLQDIAILTAGTVISEE-IGMELEKATLED-LGQAKRIVINKDNTT 331
L.pneumophila    NNQQNMSCELEHPFILLVDKKVSSIREMLSVLEGVA--KSGRPLLIIAEDVEGEALATLVVNNMRGIVKVCAVKAPGFGDRRKAMLQDIAILTKGQVISEE-IGKSLEGATLED-LGSAKRIVVTKENTT 330
M.avium hsp60    TDAERQEAVLEDPFILLVSSKVSTVKDLLPLLEKVI--QAGKPLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGGQVISEE-VGLSLESADISL-LGKARKVVVTKDETT 329
M.bovis hsp60    TDPERQEAVLEDPYILLVSSKVSTVKDLLPLLEKVI--GAGKPLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGGQVISEE-VGLTLENADLSL-LGKARKVVVTKDETT 329
M.leprae groE    TDFDSQQAVLDDPLVLLHQEKISSLPELLPMLEKVT--ESGKPLLIVAEDLEGEALATLVVNSIRKTLKAVAVKSPFFGDRRKAFLEDLAIVTGGQVVNPE-TGLVLREVGTDV-LGSARRVVVSKDDTI 329
M.leprae hsp6    TDAERQEAVLEEPYILLVSSKVSTVKDLLPLLEKVI--QAGKSLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGAQVISEE-VGLTLENTDLSL-LGKARKVVMTKDETT 329
M.tuberculosi    TDFDNQQAVLEDALILLHQDKISSLPDLLPLLEKVA--GTGKPLLIVAEDVEGEALATLVVNAIRKTLKAVAVKGPYFGDRRKAFLEDLAVVTGGQVVNPD-AGMVLREVGLEV-LGSARRVVVSKDDTV 329
N.meningitidi    NDAEKQIAGLDNPFVLLFDKKISNIRDLLPVLEQVA--KASRPLLIIAEDVEGEALATLVVNNIRGILKTVAVKAPGFGDRRKAMLQDIAILTGGTVISEE-VGLSLEKATLDD-LGQAKRIEIGKENTT 331
S.aureus hsp6    TDSDKMVAELERPYILVTDKKISSFQDILPLLEQVV--QSNRPILIVADEVEGDALTNIVLNRMRGTFTAVAVKAPGFGDRRKAMLEDLAILTGAQVITDD-LGLDLKDASIDM-LGTASKVEVTKDNTT 330
Synechocystis    TDAERMEAVLEDPRILITDKKINLVQDLVPILEQVA--RQGKPLLIIAEDIEKEALATLVVNRLRGVLNVAAVKAPGFGDRRKQMLEDIATLTGGQVISED-AGLKLESATVDS-LGSARRINITKDNTT 330
Synechocystis    TDSDRQLVEFDNPLILITDKKISAIAELVPVLEAVA--RAGRPLLIIAEDIEGEALATLVVNKARGVLNVAAIKAPAFGDRRKAVLQDIAILTGGSVISED-IGLSLDTVSLDQ-LGQAVKATLEKDNTI 329
Tsukamurella     TDAERQEAVLEDPYVLLVSGKISTVKDLLPLLEKVI--QSGKPLAIIAEDVEGEALVTLIVNKIRGTFKSVAIKAPGFGDRRKAMLQDMAILTGGQVISEE-IGLSLDTAGLEV-LGQARQVVVTKDETT 329
S.pombe hsp60   TDVKSQKVEFENPLILLSEKKVSAVQDILPSL-ELAAQQ-RRPLVIIAEDVDGEALAACILNKLRGQLQVVAIKAPGFGDNRRNMLGDLAVLTDSAVFNDE-IDVSIEKAQPHH-LGSCGSVTVTKEDTI 362
S.cerevisiae     TDPKSSKVEFEKPLLLLSEKKISSIQDILPAL-EISNQS-RRPLLIIAEDVDGEALAACILNKLRGQVKVCAVKAPGFGDNRKNTIGDIAVLTGGTVFTEE-LDLKPEQCTIEN-LGSCDSITVTKEDTV 352
P.falciparum     NN--SQKVELDKPYILIHEKKISTVKSLLPVLEHV--LQNQSSLLVIAEDVDSDALATLIVNKLRLGLKICAVKAPGFGEHRKALIHDIAVMTGAKVITEET-GLKLDDPQVVSYLGKAKSINVTKDSTL 358
Onchocerca hs    TNNEKMIVELDDPYLLITEKKLNIIQPLLSILEAVV--KSGKPLLIIAEDIEGEALSTLVINKLRGGLKVAAVKAPGFGDRRKEMLEDIAALTNAKYVIKDELGIKMEDLTLED-LGIAKNVKITKDNTT 334
C.elegans hsp    TSAKGAKVEYEKALVLLSEKKISQVQDIVPAL-ELANKL-RRPLVIIAEDVDGEALTTLVLNRLKVGLQVVAIKAPGFGDNRKNALKDMGIATGASIFGDETLDLRLEDITAND-LGEVDEVTITKDDTL 347
D.melanogaste    NSSKGAKVEFQDALLLLSEKKISSVAEHHSPLWRLASRRTRKPLVIIAEDIDGEALSTLVVNRLKIGLQVAAVKAPGFGDNRKSTLTDMA-TSGGIVFGDDVSLVKLEDVKVSD-LGQVGEVVITKDDTL 357
human hsp60      NTSKGQKCEFQDAYVLLSEKKISSIQSIVPAL-EIANAH-RKPLVIIAEDVDGEALSTLVLNRLKVGLQVVAVKAPGFGDNRKNQLKDMAIATGGAVFGEEGLTLNLEDVQPHD-LGKVGEVIVTKDDAM 356
Arabidopsis h    TNQKTQKCELDDPLILIHEKKISSINSIVKVL-ELALKR-QRPLLIVSEDVESDALATLILNKLRAGIKVCAIKAPGFGENRKANLQDLAALTGGEVITDE-LGMNLEKVDLSM-LGTCKKVTVSKDDTV 361
maize hsp60      TNSKAQKCEPEDPLILIHDRKVTNMHAVVKVL-EMALKK-QRPLLIVAEDVESEALGTLIINKLRAGIKVCAVKAPGFGENRKANLQDLAILTGGEVITEE-LGMNLENVEPHM-LGSCKKVTVSKDDTV 364
RUBISCO chape    TNPEKSIVEFENARVLITDQKISAIKDIIPLLEKTT--QLRAPLLIISEDITGEALATLVVNKLRGILNVAAIKAPGFGERRKALLQDIAILTGAEFQASD-LGLLVENTTIEQ-LGLARKVTISKDSTT 376
```

Fig. 10C

IVDGAGD--AAIAGRVAQIRSQIEESI-SDYDKEKLQERLAKLAGGVAVIKVGAATEVELKERKDRVEDALNATRAAVEEGIVPGGGVALLRAAPALDKLKTE--NGDEATGVNIVLRALEAPLRQIAE
400 410 420 430 440 450 460 470 480 490 500 510 520
S.pneumoniae IIDGVGD--EAAIQGRVTQIRQQIEEAT-SDYDREKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVALIRVASKIAGLKGQ--NEDQNVGIKVALRAMESPLRQIVL 456
S.pyogenes hs IIDGVGD--EAAIQGRVGQIRKQIEEAT-SDYDREKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVDDALHATRAAVEEGVVAGGGVALVRVAAKLSGLTAQ--NEDQNVGIKVALRAMEAPLRQIVS 456
S.pneumoniae IVEGAGN--PEAISHRVAVIKSQIETTT-SEFDREKLQERLAKLSGGVAVIKVGAATETELKEMKLRIEDALNATRAAVEEGIVAGGGTALANVIPAVATLELT---GDEATGRNIVLRALEEPVRQIAH 453
S.pyogenes hs IVEGSGS--SEAIANRIALIKSQLETTT-SDFDREKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITVIEKVAALELE---GDDATGRNIVLRALEEPVRQIAL 453
B.subtilis gr IVEGAGE--TDKISARVTQIRAQVEETT-SEFDREKLQERLAKLAGGVAVIKVGAATETELKERKLRIEDALNSTRAAVEEGIVSGGGTALVNVYNKVAAVEAE---GDAQTGINIVLRALEEPIRQIAH 453
Clostridium h IVNGRGN--SEEIKNRINQIKLQLEATT-SEFDKEKLQERLAKLAGGVAVVKVGAATETELKESKLRIEDALAATKAAVEEGIVPGGGTAYVNVINEVAKLTSD--IQDEQVGINIIVRSLEEPMRQIAH 454
Cowdria hsp60 II-GSVDNSCAHVQSRICQIRMQIDNST-SDYDKEKLQERLAKLSGGVAVLKVGGSSEVEVKERKDRVEDALHATRAAVEEGVVPGGGAALLYTLSALDNLKSK--NDDEQLGINIYKRALQAPIKRIIK 458
Haemophilus h IIDGIGD--EAQIQGRVAQIRQQIEEST-SDYDKEKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVADALHATRAAVEEGIVAGGGVALIRAAGRVVGLQGE--NEEQNVGIKLALRAMEAPLRQIVA 456(8)
L.pneumophila IIDGEGK--ATEINARITQIRAQMEETT-SDYDREKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGIVAGGGVALIRAQKALDSLKGD--NDDQNMGINILRRAIESPMRQIVT 455
M.avium hsp60 IVEGAGD--SDAIAGRVAQIRTEIENSD-SDYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVALLHAIPALDELKPE---GEEATGANIVRVALERPLKQIAF 453
M.bovis hsp60 IVEGAGD--TDAIAGRVAQIRQEIENSD-SDYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVTLLQAAPTLDELKLE---GDEATGANIVKVALEAPLKQIAF 453
M.leprae groE IVDGGG--SNDAVAKRVNQLRAEIEVSD-SEWDREKLQERVAKLAGGVAVIKVGAVTETALKKRKESVEDAVAAAKASIEEGIIAGGGSALVQCGAALKQLRTSLTGDEALGIDVFFEALKAPLYWIAT 455
M.leprae hsp6 IVEGAGD--TDAIAGRVAQITEIENSD-SDYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVTLLQAAPALDKLKLT---GDEATGANIVKVALEAPLKQIAF 453
M.tuberculosi IVDGGG--TAEAVANRAKHLRAEIDKSD-SDWDREKLGERLAKLAGGVAVIKVGAATETALKERKESVEDAVAAAKAAVEEGIVPGGGASLIHQARKALTELRASLTGDEVLGVDVFSEALAAPLFWIAA 456
N.meningitidi IIDGFGD--AAQIEARVAEIRQQIETAT-SDYDKEKLQERVAKLAGGVAVIKVGAATEVEMKEKKDRVEDALHATRAAVEEGVVAGGGVALLRARAALENLHTG--NADQDAGVQIVLRAVESPLRQIVA 456
S.aureus hsp60 VVDGDGD--ENSIDARVSQLKSQIEETE-SDFDREKLQERLAKLAGGVAVIKVGAASETELKERKLRIEDALNSTRAAVEEGIVAGGGTALVNVYQKVSENEAE---GDIETGVNIVLKALTAPVRQIAE 454
Synechocystis IVAEGNE---AAVKSRCEQIRRQIEETDS-SYDKEKLQERLAKLAGGVAVIKVGAATETEMKDRKLRLEDAINATKAAVEEGIVPGGGTTLAHLAPQLEDWATGNLKQEELTGALIVARALPAPLKRIAE 456
Synechocystis LVAGADKRASAGVKERIEQLRKEYAASD-SDYDKEKIQERIAKLAGGVAVIKVGAATETELKDRKLRIEDALNATKAAVEEGIVPGGGTTLIRLAGKIESFKAQLSNDEERVAADIIAKALEAPLHQLAS 458
Tsukamurella IVDGAGS--KEQIAGRVSQIRAEIENSD-SDYDREKLQERLAKLAGGVAVIKAGAATE-DLKERKHRIEDAVRNAKAAVEEGIVAGGGSSLAQSGTVFDSYALE---GDEATGANIVKVALDAPVKQIAV 452
S.pombe hsp60 IMKGAGDH--VKVNDRCEQIRGVMADPNLTESEKEKLQERLAKLSGGIAVIKVGASSEVEVNEKKDRIVDALNAVKAAVSEGVLPGAGTSFVKASLRLGDIPTN--NFDQKLGVEIVRKAITRPAQTILE 488
S.cerevisiae ILNGSGPK--EAIQERIEQIKGSIDITTTNSYEKEKLQERLAKLSGGVAVIRVGGASEVEVGEKKDRYDDALNATRAAVEEGILPGGGTALVKASRVLDEVVVD--NFDQKLGVDIIRKAITRPAKQIIE 478
P.falciparum IMEGEGKK--EEINERCESIRNAIKNNT-SDYEKEKLQERLAKITGGVALIKVGGISEVEVNEIKDRIQDALCATKAAVEEGIVPGGGSALLFASKELDSVQTD--NYDQRVGVNIIKDACKAPIKQIAE 483
Onchocerca hs IV-SE-NRVTDRVKARIEQIKSQIESST-SDYDKEKLRERLAKLSGGVAVLKVGGATELEVKERRDRVEDQLHATRAAIEEGIVPGGGVALLYASSALDKLKGA--DDEEQIGINIIKKVLSVPIKRLVK 459
C.elegans hsp LLRGRGDQ--TEIEKRIEEITDEIERST-SDYEKEKLNERLAKLSKGVAVLKIGGGSEVEVGEKKDRVTDALCATRAAVEEGIVPGGGVALLRSLTALKNYKAA--NEDQQIGVNIVKKALTQPIATIVK 472
D.melanogaste LLKGKGKK--DDVLRRANQIRTKIEDTT-SEYEKEKLQERLARLASGVA-LRVGGSSEVEVNEKKDRVHDALNATRAAVEEGIVLGGGCALLRCIPALDSLTPA--NEDQKIGIEIIKRTLKIPAMTIAK 481
human hsp60 LLKGKGDK--AQIEKRIQEIIEQLDVTT-SEYEKEKLNERLAKLSDGVAVLKVGGTSDVEVNEKKDRVTDALNATRAAVEEGIVLGGGCALLRCIPALDSLTPA--NEDQKIGIEIIKRTLKIPAMTIAK 481
Arabidopsis h ILDGAGDK--KGIEERCEQIRSAIELST-SDYDKEKLQERLAKLSGGVAVLKIGGASEAEVGEKKDRVTDALNATKAAVEEGILPGGGVALLYAARELEKLPTA--NFDQKIGVQIIQNALKTPVYTIAS 486
maize hsp60 ILDGAGDK--KSIEERADQIRSAVENST-SDYDKEKLQERLAKLSGGVAVLKIGGASEAEVGEKKDRVTDALNATKAAVEEGIVPGGGVALLYASKELDKLQTA--NFDQKIGVQIIQNALKTPVHTIAS 489
RUBISCO chape IIADAASK--DELQSRVAQLKKELSETD-SIYDSEKLAERIAKLSGGVAVIKVGAATETELEDRKLRIEDAKNATFAAIEEGIVPGGGTALVHLSGYVPAIKEKLEDADERLGADIVQKALVAPAALIAQ 503

Fig. 10D

NAGLEGSV-VVEKVKN---SEAG-GYNAATGEYVDMIAAGIIDPTKVTRSALQNAASVASLMLTTEAVVVDKPEKEAAPAG-MPGM--MGGMGGMGGM--M--

530 540 550 560 570 580 590 600 610 620

S.pneumoniae	NCGEEPSV-VANTVKA---GDGNYGYNAATEEYGNMIDMGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKGDAPDLG-AAG--GMGGMGG	543 (MM)
S.pyogenes hs	NAGEEPSV-VTNNVKA---GEGNYGYNAATEEYGNMIDFGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKGDAPDLG-AAG---MGGMGGM------M	544
S.pneumoniae	NAGFEGSI-VIDRLKN---AELGIGFNAATGEWVNMIDQGIIDPVKVSRSALQNAASVASLILTTEAVVANKPEPVAPAPA-M----DPSMMGGMGG	541
S.pyogenes hs	NAGYEGSV-VIDKLKN---SPAGTGFNAATGEWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPA-MPAGMDPGMMGGMGG	545
B.subtilis gr	NAGLEGSV-IVERLKN---EEIGVGFNAATGEWVNMIEKGIVDPTKVTRSALQNAASVAAMFLTTEAVVADKPEENGGGAG-MP---DMGGMGGMGGM--M	544
Clostridium h	NAGLEGSV-IIEKVKN---SDAGVGFDALRGEYKDMIKAGIVDPTKVTRSALQNAASVASTFLTTEAAVADIPEK-----E-MPQGAGM----GMDGM--Y	539
Cowdria hsp60	NAGSENAPCVIAHLLKQNDKELI--FNVDVTNFANAFTSGVIDPLKVVRIAFDFAVSLAAVFMTLNAIVVDIPSKDDNSAAGGAGMGGMGGMGG--------F	551
Haemophilus h	NSGEEASV-IASAVKN---GEGNFGYNAGTEQYGDMIAMGILDPTKVTRSALQFAASVAGLMITTECMVTELPKDDKADLG-AAGMGGMGGMGGM-------M	547(8)
L.pneumophila	NAGYEASV-VVNKVAE---HKDNYGFNAATGEYGDMVEMGILDPTKVTRMALQNAASVASLMLTTECMVADLPKKEEG-VG-AGDMGGMGGMGGMGGM--MZ	549
M.avium hsp60	NGGLEPGV-VAEKVRN---SPAGTGLNAATGEYEDLLKAGIADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKAAAPAG-DPTG----GMGGMD----F	541
M.bovis hsp60	NSGLEPGV-VAEKVRN---LPAGHGLNAQTGVYEDLLAAGVADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKEKASVP-G-GG----DMGGMD----F	540
M.leprae groE	NAGLDGAV-VVDKVSG---LPAGHGLNASTLGYGDLVADGVVDPVKVTRSAVLNAASVARMMLTTETAVVDKPAKTEEHDH-H-GHAH	537
M.leprae hsp6	NSGMEPGV-VAEKVRN---LSVGHGLNAATGEYEDLLKAGVADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKTAAPAS-DPTG----GMGGMD----F	541
M.tuberculosi	NAGLDGSV-VVKKVSE---LPAGHGLNVNTLSYGDLAADGVIDPVKVTRSAVLNAASSVARMVLTTETVVVDKPAKAEDHDH-HHGHAH	539
N.meningitidi	NAGGEPSV-VVNKVLE---GKGNYGYNAGSGEYGDMIEMGVLDPAKVTRSALQHAASIAGLMLTTDCMIAEIPEEKPAMPD-MGGMGGMGGM--M------Z	545
S.aureus hsp6	NAGLEGSV-IVERLKN---AEPGVGFNGATNEWVNMLRRGIVDPTKVTRSALQHAASVAAMFLTTEAVVASIPEKNNDQPN-M---------GGMPGM--M	539
Synechocystis	NAGQNGAV-ISERVKE---KEFNVGYNAASLEYVDMLAAGIVDPAKVTRSALQNAASIAGMVLTTECIVVDKPEKEKAP---A---------GAPGGD--FDY	541
Synechocystis	NAGVEGSV-IVEKVKE---ATGNQGYNVITGKIEDLIAAGIIDPAKVVRSALQNAASIAGMVLTTEALVVEKPEPAAPA---MPDMGGMGGMGGMGGM--GMM	552
Tsukamurella	NAGLEPGV-VAEKVRN---SPAGTGLNAATGVYEDLLAAGINDPVKVTRSALQNAASIAALFLTTEAVVADKPEKAGAPV--DPTG----GMGGMD----F	539
S.pombe hsp60	NAGLEGNL-IVGKLKELYGKEFNIGYDIAKDRFVDLNEIGVLDPLKVVRTGLVDASGVASLMGTTECAIVDAPEESKAPAGP-PGM------GGMGGMPGMM	583
S.cerevisiae	NAGEEGSV-IIGKLIDEYGDDFAKGYDASKSEYTDMLATGIIDPFKVVRSGLVDASGVASLLATTEVAIVDAPE-PPAAAGA-GGMP-----GGMPGMPGMM	573
P.falciparum	NAGHEGSV-VAGNILKDKNSNI--GFNAQEGKYVDMIESGIIDPTKVVKTAISDAASIASLMTTTEVAIVDFKDSKNEESSQ--HMNSVNSMGDMGGM----Y	577
Onchocerca hs	NAGLESAV-IIDYLIKQNNKELI--YNVEAMSYANAFAAGVIDPAKVVRIAFETAISVASVLITTESMIVDIPNKDEN-ASSPMGAGGMGRMND--------F	550
C.elegans hsp	NAGLEPSS-IIDEVTGNSNTSY--GYDALNGKFVDMFEAGIIDPTKVVRTALQDASGVASLLATTECVVTEIPKEEAVGGPA-GGMGGMGGMGG---MGGMGF	568
D.melanogaste	NAGVDGAM-VVAKVENQAG-DY--GYDA-KGEYGNLIEKGIIDPTKVVRTAITDASGVASLLTTAEAVVTEIPKED--GAPAMPGMGGMGGMGGMGGMGGMM	576
human hsp60	NAGVEGSL-IVEKIM-QSSSEV--GYDAMAGDFVNMVEKGIIDPTKVVRTALLDAAGVASLLTTAEVVVTEIPKEEK----D-PGMGAMGGMGG--GMGGGMF	573
Arabidopsis h	NAGVEGAV-IVGKLLEQDNPDL--GYDAAKGEYVDMVKAGIIDPLKVIRTALVDAASVSSLLTTTEAVVVDLPKDESESGAA-GG--------GMGGMVVMDY	577
maize hsp60	NAGVEGAV-VVGKLLEQGNTDL--GYDAAKDEYVDMVKAGIIDPLKVIRTALVDAASVSSLMTTTESIIVEIPKEE----AP-AP--------AMGGMGGMDY	576
RUBISCO chape	NAGIEGEV-VVEKIKN---GEWEVGYNAMTDTYENLVESGVIDPAKVTRCALQNAASVAGMVLTTQAIVVEKPKPKAAVAA-APQG-------------LTI	587

Fig. 10E

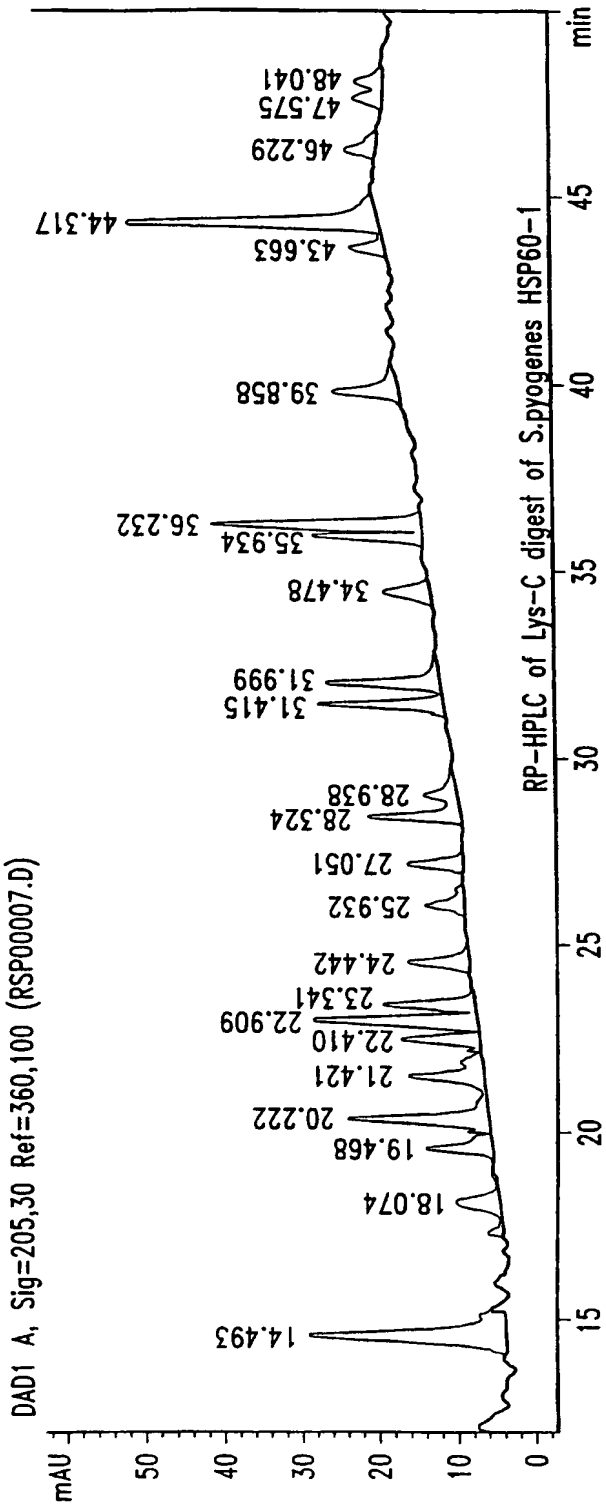

DAD1 A, Sig=205,30 Ref=360,100 (RSP00007.D)
mAU
0
10
20
30
40
50
RP-HPLC of Lys-C digest of S.pyogenes HSP60-1
min
15
20
25
30
35
40
45
14.493
18.074
19.468
20.222
21.421
22.410
22.909
23.341
24.442
25.932
27.051
28.324
28.938
31.415
31.999
34.478
35.934
36.232
39.858
43.663
44.317
46.229
47.575
48.041

Fig. 11A

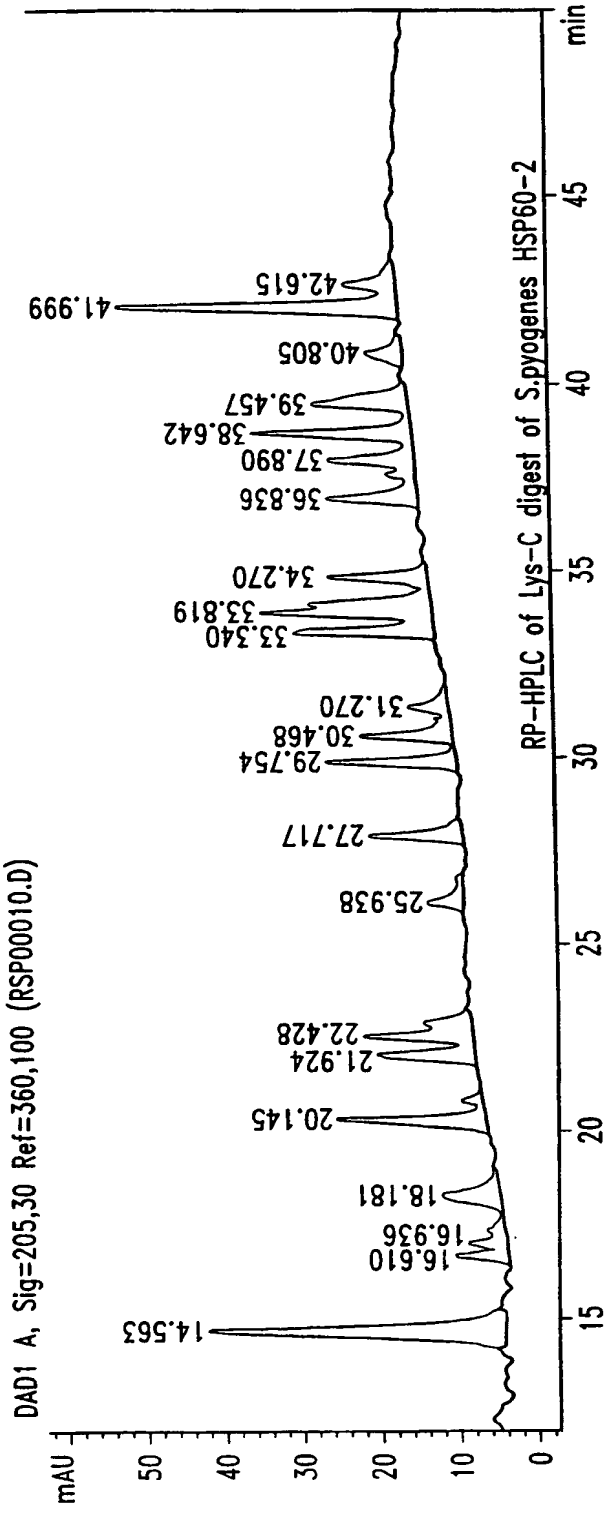

RP-HPLC of Lys-C digest of S.pyogenes HSP60-2
DAD1 A, Sig=205,30 Ref=360,100 (RSP00010.D)
mAU
min
14.563
16.610
16.936
18.181
20.145
21.924
22.428
25.938
27.717
29.754
30.468
31.270
33.340
33.819
34.270
36.836
37.890
38.642
39.457
40.805
41.999
42.615

Fig. 11B

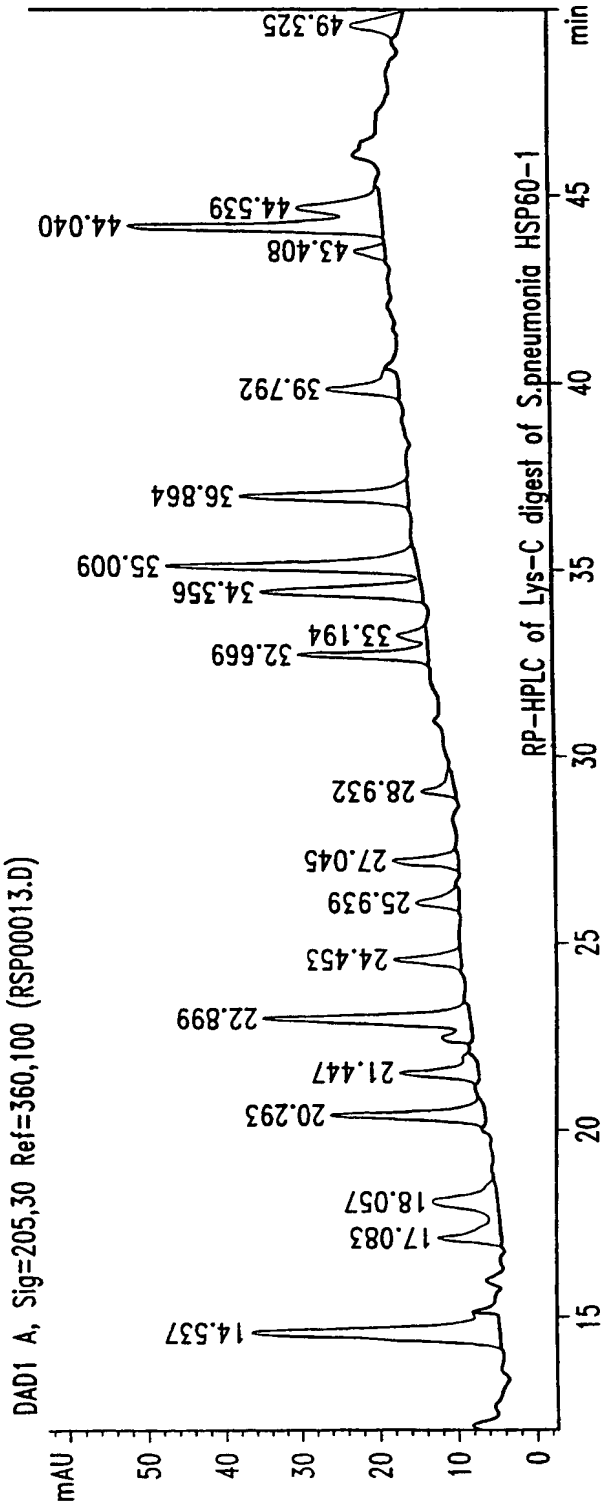

RP-HPLC of Lys-C digest of S.pneumonia HSP60-1
DAD1 A, Sig=205,30 Ref=360,100 (RSP00013.D)
mAU
min
14.537
17.083
18.057
20.293
21.447
22.899
24.453
25.939
27.045
28.932
32.669
33.194
34.356
35.009
36.864
39.792
43.408
44.040
44.539
49.325

Fig. 11C

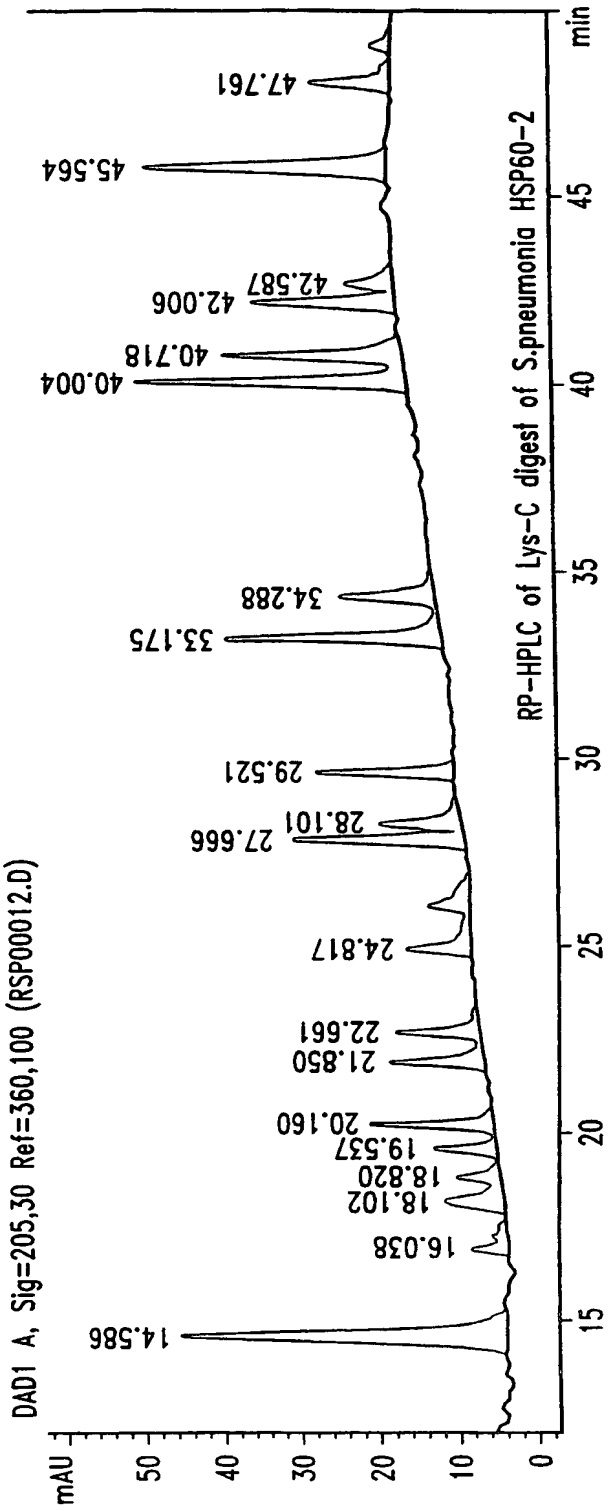

RP-HPLC of Lys-C digest of S.pneumonia HSP60-2
DAD1 A, Sig=205,30 Ref=360,100 (RSP00012.D)
mAU
min
14.586
16.038
18.102
18.820
19.537
20.160
21.850
22.661
24.817
27.666
28.101
29.521
33.175
34.288
40.004
40.718
42.006
42.587
45.564
47.761

Fig. 11D

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Stessgen Biotechnologies Corporation

(ii) TITLE OF INVENTION: STREPTOCOCCAL HEAT SHOCK PROTEINS OF THE HSP60 FAMILY

(iii) NUMBER OF SEQUENCES: 91

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Gowling, Strathy & Henderson
(B) STREET: Commerce Court West, Suite 4900
(C) CITY: Toronto
(D) STATE: Ontario
(E) COUNTRY: Canada
(F) ZIP: M5L 1J3

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 29 December 1998
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Omar A. Nassif
(B) REGISTRATION NUMBER: 4016
(C) REFERENCE/DOCKET NUMBER: T8464440WO

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (416) 862-7525
(B) TELEFAX: (416) 862-7661

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1665 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 15..1649

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GAATTCGGCT TCAT ATG GCG GCT AAA GAC GTA AAA TTC GGT AAC GAC GCT          50
                Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala
                  1               5                  10

CGT GTG AAA ATG CTG CGC GGC GTA AAC GTA CTG GCA GAT GCA GTG AAA          98
Arg Val Lys Met Leu Arg Gly Val Asn Val Leu Ala Asp Ala Val Lys
          15                  20                  25

GTT ACC CTC GGC CCA AAA GGC CGT AAC GTA GTT CTG GAT AAA TCT TTC         146
Val Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe
      30                  35                  40

GGT GCA CCG ACC ATC ACT AAA GAT GGT GTT TCC GTA GCA CGT GAA ATC         194
Gly Ala Pro Thr Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile
  45                  50                  55                  60

GAA CTG GAA GAC AAG TTC GAA AAC ATG GGT GCG CAG ATG GTG AAA GAA         242
Glu Leu Glu Asp Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu
                  65                  70                  75

GTT GCC TCT AAA GCG AAC GAC GCT GCA GGT GAC GGT ACC ACC ACC GCA         290
Val Ala Ser Lys Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala
              80                  85                  90

ACC GTA CTG GCT CAG TCC ATC ATC ACT GAA GGC CTG AAA GCC GTT GCT         338
Thr Val Leu Ala Gln Ser Ile Ile Thr Glu Gly Leu Lys Ala Val Ala
          95                  100                 105

GCG GGC ATG AAC CCG ATG GAT CTG AAA CGT GGT ATC GAC AAA GCT GTC         386
Ala Gly Met Asn Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val
      110                 115                 120

GCT GCT GCT GTT GAA GAA CTG AAA GCA CTG TCC GTA CCG TGC TCC GAC         434
Ala Ala Ala Val Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp
  125                 130                 135                 140

TCT AAA GCT ATT GCT CAG GTT GGT ACC ATC TCC GCT AAC TCC GAC GAA         482
Ser Lys Ala Ile Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Glu
                  145                 150                 155

ACC GTA GGT AAA CTG ATC GCT GAA GCG ATG GAC AAA GTC GGT AAA GAA         530
Thr Val Gly Lys Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys Glu
              160                 165                 170

GGC GTG ATC ACC GTT GAA GAC GGT ACC GGT CTG CAG GAC GAA CTG GAC         578
Gly Val Ile Thr Val Glu Asp Gly Thr Gly Leu Gln Asp Glu Leu Asp
          175                 180                 185

GTG GTT GAA GGT ATG CAG TTC GAC CGT GGC TAC CTG TCT CCT TAC TTC         626
Val Val Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe
      190                 195                 200
```

```
ATC AAC AAG CCG GAA ACT GGC GCA GTA GAA TTG GAA AGC CCG TTC ATC        674
Ile Asn Lys Pro Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile
205                 210                 215                 220

CTG CTG GCT GAC AAG AAA ATC TCC AAC ATC CGC GAA ATG CTG CCG GTT        722
Leu Leu Ala Asp Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val
                225                 230                 235

CTG GAA GCT GTA GCG AAA GCA GGC AAA CCG CTG CTG ATC ATC GCT GAA        770
Leu Glu Ala Val Ala Lys Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu
            240                 245                 250

GAT GTT GAA GGC GAA GCG CTG GCA ACT CTG GTT GTT AAC ACC ATG CGC        818
Asp Val Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg
        255                 260                 265

GGT ATC GTA AAA GTC GCT GCG GTT AAA GCA CCT GGC TTC GGC GAT CGT        866
Gly Ile Val Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg
    270                 275                 280

CGT AAA GCA ATG CTG CAG GAT ATC GCT ACC CTG ACC GGT GGT ACC GTT        914
Arg Lys Ala Met Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val
285                 290                 295                 300

ATC TCT GAA GAG ATC GGT ATG GAG CTG GAA AAA GCA ACT CTG GAA GAT        962
Ile Ser Glu Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp
                305                 310                 315

CTG GGC CAG GCG AAA CGC GTT GTT ATC AAC AAA GAT ACC ACC ACC ATC       1010
Leu Gly Gln Ala Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile
            320                 325                 330

ATC GAT GGC GTG GGC GAC GAA GCT GCA ATC CAG GGT CGC GTG ACT CAG       1058
Ile Asp Gly Val Gly Asp Glu Ala Ala Ile Gln Gly Arg Val Thr Gln
        335                 340                 345

ATT CGT CAG CAG ATC GAA GAA GCA ACT TCC GAC TAT GAC CGT GAA AAA       1106
Ile Arg Gln Gln Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys
    350                 355                 360

CTG CAG GAG CGC GTA GCG AAA CTG GCA GGC GGC GTT GCG GTT ATC AAA       1154
Leu Gln Glu Arg Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys
365                 370                 375                 380

GTT GGT GCT GCG ACT GAA GTT GAA ATG AAA GAG AAG AAA GCC CGC GTT       1202
Val Gly Ala Ala Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val
                385                 390                 395

GAA GAT GCC CTG CAC GCT ACC CGT GCT GCG GTA GAA GAA GGC GTG GTT       1250
Glu Asp Ala Leu His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val
            400                 405                 410

GCT GGT GGT GGC GTT GCG CTG ATT CGC GTA GCG TCT AAA ATT GCC GGC       1298
Ala Gly Gly Gly Val Ala Leu Ile Arg Val Ala Ser Lys Ile Ala Gly
        415                 420                 425
```

```
CTG AAA GGT CAG AAC GAA GAC CAG AAC GTA GGT ATC AAA GTT GCG CTG      1346
Leu Lys Gly Gln Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu
    430                 435                 440

CGC GCA ATG GAA TCC CCA CTG CGT CAA ATC GTA CTG AAC TGC GGC GAA      1394
Arg Ala Met Glu Ser Pro Leu Arg Gln Ile Val Leu Asn Cys Gly Glu
445                 450                 455                 460

GAG CCG TCT GTA GTG GCT AAC ACC GTG AAA GCC GGT GAC GGT AAC TAC      1442
Glu Pro Ser Val Val Ala Asn Thr Val Lys Ala Gly Asp Gly Asn Tyr
                465                 470                 475

GGT TAC AAC GCT GCA ACT GAA GAA TAC GGC AAC ATG ATC GAC ATG GGT      1490
Gly Tyr Asn Ala Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Met Gly
            480                 485                 490

ATC CTG GAT CCA ACC AAA GTA ACT CGT TCT GCT CTG CAG TAC GCG GCT      1538
Ile Leu Asp Pro Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala
        495                 500                 505

TCT GTT GCG GGT CTG ATG ATC ACC ACC GAG TGC ATG GTT ACC GAC CTG      1586
Ser Val Ala Gly Leu Met Ile Thr Thr Glu Cys Met Val Thr Asp Leu
    510                 515                 520

CCG AAA GGC GAT GCA CCT GAC TTA GGT GCT GCT GGT GGT ATG GGC GGC      1634
Pro Lys Gly Asp Ala Pro Asp Leu Gly Ala Ala Gly Gly Met Gly Gly
525                 530                 535                 540

ATG GGC GGA ATG ATG TGATCAAGCC GAATTC                                1665
Met Gly Gly Met Met
                545
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 545 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met
  1               5                  10                  15

Leu Arg Gly Val Asn Val Leu Ala Asp Ala Val Lys Val Thr Leu Gly
            20                  25                  30

Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala Pro Thr
        35                  40                  45

Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile Glu Leu Glu Asp
```

```
     50                  55                  60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys
 65                  70                  75                  80

Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
                 85                  90                  95

Gln Ser Ile Ile Thr Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn
            100                 105                 110

Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Ala Ala Ala Val
        115                 120                 125

Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp Ser Lys Ala Ile
    130                 135                 140

Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Glu Thr Val Gly Lys
145                 150                 155                 160

Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys Glu Gly Val Ile Thr
                165                 170                 175

Val Glu Asp Gly Thr Gly Leu Gln Asp Glu Leu Asp Val Val Glu Gly
            180                 185                 190

Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Lys Pro
        195                 200                 205

Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu Leu Ala Asp
    210                 215                 220

Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu Glu Ala Val
225                 230                 235                 240

Ala Lys Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly
                245                 250                 255

Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile Val Lys
            260                 265                 270

Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met
        275                 280                 285

Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val Ile Ser Glu Glu
    290                 295                 300

Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly Gln Ala
305                 310                 315                 320

Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp Gly Val
                325                 330                 335

Gly Asp Glu Ala Ala Ile Gln Gly Arg Val Thr Gln Ile Arg Gln Gln
            340                 345                 350
```

```
Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg
        355                 360                 365

Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala
    370                 375                 380

Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val Glu Asp Ala Leu
385                 390                 395                 400

His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly
                405                 410                 415

Val Ala Leu Ile Arg Val Ala Ser Lys Ile Ala Gly Leu Lys Gly Gln
            420                 425                 430

Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu Arg Ala Met Glu
        435                 440                 445

Ser Pro Leu Arg Gln Ile Val Leu Asn Cys Gly Glu Glu Pro Ser Val
    450                 455                 460

Val Ala Asn Thr Val Lys Ala Gly Asp Gly Asn Tyr Gly Tyr Asn Ala
465                 470                 475                 480

Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Met Gly Ile Leu Asp Pro
                485                 490                 495

Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala Ser Val Ala Gly
            500                 505                 510

Leu Met Ile Thr Thr Glu Cys Met Val Thr Asp Leu Pro Lys Gly Asp
        515                 520                 525

Ala Pro Asp Leu Gly Ala Ala Gly Gly Met Gly Gly Met Gly Gly Met
    530                 535                 540

Met
545
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1654 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 15..1637

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GAATTCGGCT TCAT ATG GCA AAA GAA ATT AAA TTT TCA TCA GAT GCC CGT          50
                Met Ala Lys Glu Ile Lys Phe Ser Ser Asp Ala Arg
                  1               5                  10

TCA GCT ATG GTC CGT GGT GTC GAT ATC CTT GCA GAT ACT GTT AAA GTA          98
Ser Ala Met Val Arg Gly Val Asp Ile Leu Ala Asp Thr Val Lys Val
          15                  20                  25

ACT TTG GGA CCA AAA GGT CGC AAT GTC GTT CTT GAA AAG TCA TTC GGT         146
Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Glu Lys Ser Phe Gly
      30                  35                  40

TCA CCC TTG ATT ACC AAT GAC GGT GTG ACT ATT GCC AAA GAA ATT GAA         194
Ser Pro Leu Ile Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu
 45                  50                  55                  60

TTA GAA GAC CAT TTT GAA AAT ATG GGT GCC AAA TTG GTA TCA GAA GTA         242
Leu Glu Asp His Phe Glu Asn Met Gly Ala Lys Leu Val Ser Glu Val
                  65                  70                  75

GCT TCA AAA ACC AAT GAT ATC GCA GGT GAT GGA ACT ACA ACT GCA ACT         290
Ala Ser Lys Thr Asn Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr
              80                  85                  90

GTT TTG ACC CAA GCA ATC GTC CGT GAA GGA ATC AAA AAC GTC ACA GCA         338
Val Leu Thr Gln Ala Ile Val Arg Glu Gly Ile Lys Asn Val Thr Ala
          95                  100                 105

GGT GCA AAT CCA ATC GGT ATT CGT CGT GGG ATT GAA ACA GCA GTT GCC         386
Gly Ala Asn Pro Ile Gly Ile Arg Arg Gly Ile Glu Thr Ala Val Ala
     110                 115                 120

GCA GCA GTT GAA GCT TTG AAA AAC AAC GTC ATC CCT GTT GCC AAT AAA         434
Ala Ala Val Glu Ala Leu Lys Asn Asn Val Ile Pro Val Ala Asn Lys
125                 130                 135                 140

GAA GCT ATC GCT CAA GTT GCA GCC GTA TCT TCT CGT TCT GAA AAA GTT         482
Glu Ala Ile Ala Gln Val Ala Ala Val Ser Ser Arg Ser Glu Lys Val
                  145                 150                 155

GGT GAG TAC ATC TCT GAA GCA ATG GAA AAA GTT GGC AAA GAC GGT GTC         530
Gly Glu Tyr Ile Ser Glu Ala Met Glu Lys Val Gly Lys Asp Gly Val
              160                 165                 170

ATC ACC ATC GAA GAG TCA CGT GGT ATG GAA ACA GAG CTT GAA GTC GTA         578
Ile Thr Ile Glu Glu Ser Arg Gly Met Glu Thr Glu Leu Glu Val Val
          175                 180                 185

GAA GGA ATG CAG TTT GAC CGT GGT TAC CTT TCA CAG TAC ATG GTG ACA         626
Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Gln Tyr Met Val Thr
     190                 195                 200

GAT AGC GAA AAA ATG GTG GCT GAC CTT GAA AAT CCG TAC ATT TTG ATT         674
```

```
Asp Ser Glu Lys Met Val Ala Asp Leu Glu Asn Pro Tyr Ile Leu Ile
205                 210                 215                 220

ACA GAC AAG AAA ATT TCC AAT ATC CAA GAA ATC TTG CCA CTT TTG GAA      722
Thr Asp Lys Lys Ile Ser Asn Ile Gln Glu Ile Leu Pro Leu Leu Glu
                225                 230                 235

AGC ATT CTC CAA AGC AAT CGT CCA CTC TTG ATT ATT GCG GAT GAT GTG      770
Ser Ile Leu Gln Ser Asn Arg Pro Leu Leu Ile Ile Ala Asp Asp Val
            240                 245                 250

GAT GGT GAG GCT CTT CCA ACT CTT GTT TTG AAC AAG ATT CGT GGA ACC      818
Asp Gly Glu Ala Leu Pro Thr Leu Val Leu Asn Lys Ile Arg Gly Thr
        255                 260                 265

TTC AAC GTA GTA GCA GTC AAG GCA CCT GGT TTT GGT GAC CGT CGC AAA      866
Phe Asn Val Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys
    270                 275                 280

GCC ATG CTT GAA GAT ATC GCC ATC TTA ACA GGC GGA ACA GTT ATC ACA      914
Ala Met Leu Glu Asp Ile Ala Ile Leu Thr Gly Gly Thr Val Ile Thr
285                 290                 295                 300

GAA GAC CTT GGT CTT GAG TTG AAA GAT GCG ACA ATT GAA GCT CTT GGT      962
Glu Asp Leu Gly Leu Glu Leu Lys Asp Ala Thr Ile Glu Ala Leu Gly
                305                 310                 315

CAA GCA GCG AGA GTG ACC GTG GAC AAA GAT AGC ACG GTT ATT GTA GAA     1010
Gln Ala Ala Arg Val Thr Val Asp Lys Asp Ser Thr Val Ile Val Glu
            320                 325                 330

GGT GCA GGA AAT CCT GAA GCG ATT TCT CAC CGT GTT GCG GTT ATC AAG     1058
Gly Ala Gly Asn Pro Glu Ala Ile Ser His Arg Val Ala Val Ile Lys
        335                 340                 345

TCT CAA ATC GAA ACT ACA ACT TCT GAA TTT GAC CGT GAA AAA TTG CAA     1106
Ser Gln Ile Glu Thr Thr Thr Ser Glu Phe Asp Arg Glu Lys Leu Gln
    350                 355                 360

GAA CGC TTG GCC AAA TTG TCA GGT GGT GTA GCG GTT ATT AAG GTC GGA     1154
Glu Arg Leu Ala Lys Leu Ser Gly Gly Val Ala Val Ile Lys Val Gly
365                 370                 375                 380

GCC GCA ACT GAA ACT GAG TTG AAA GAA ATG AAA CTC CGC ATT GAA GAT     1202
Ala Ala Thr Glu Thr Glu Leu Lys Glu Met Lys Leu Arg Ile Glu Asp
                385                 390                 395

GCC CTC AAC GCT ACT CGT GCA GCT GTT GAA GAA GGT ATT GTT GCA GGT     1250
Ala Leu Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly
            400                 405                 410

GGT GGA ACA GCT CTT GCC AAT GTG ATT CCA GCT GTT GCT ACC TTG GAA     1298
Gly Gly Thr Ala Leu Ala Asn Val Ile Pro Ala Val Ala Thr Leu Glu
        415                 420                 425
```

```
TTG ACA GGA GAT GAA GCA ACA GGA CGT AAT ATT GTT CTC CGT GCT TTG      1346
Leu Thr Gly Asp Glu Ala Thr Gly Arg Asn Ile Val Leu Arg Ala Leu
    430                 435                 440

GAA GAA CCT GTT CGT CAA ATT GCT CAC AAT GCA GGA TTT GAA GGA TCT      1394
Glu Glu Pro Val Arg Gln Ile Ala His Asn Ala Gly Phe Glu Gly Ser
445                 450                 455                 460

ATC GTT ATC GAT CGT TTG AAA AAT GCT GAG CTT GGT ATA GGA TTC AAC      1442
Ile Val Ile Asp Arg Leu Lys Asn Ala Glu Leu Gly Ile Gly Phe Asn
                465                 470                 475

GCA GCA ACT GGC GAG TGG GTT AAC ATG ATT GAT CAA GGT ATC ATT GAT      1490
Ala Ala Thr Gly Glu Trp Val Asn Met Ile Asp Gln Gly Ile Ile Asp
            480                 485                 490

CCA GTT AAA GTG AGT CGT TCA GCC CTA CAA AAT GCA GCA TCT GTA GCC      1538
Pro Val Lys Val Ser Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala
        495                 500                 505

AGC TTG ATT TTG ACA ACA GAA GCA GTC GTA GCC AAT AAA CCA GAA CCA      1586
Ser Leu Ile Leu Thr Thr Glu Ala Val Val Ala Asn Lys Pro Glu Pro
    510                 515                 520

GTA GCC CCA GCT CCA GCA ATG GAT CCA AGT ATG ATG GGT GGA ATG GGC      1634
Val Ala Pro Ala Pro Ala Met Asp Pro Ser Met Met Gly Gly Met Gly
525                 530                 535                 540

GGA TGATCAAAGC CGAATTC                                              1654
Gly
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 541 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Ala Lys Glu Ile Lys Phe Ser Ser Asp Ala Arg Ser Ala Met Val
  1               5                  10                  15

Arg Gly Val Asp Ile Leu Ala Asp Thr Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Ser Phe Gly Ser Pro Leu Ile
        35                  40                  45

Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp His
    50                  55                  60
```

```
Phe Glu Asn Met Gly Ala Lys Leu Val Ser Glu Val Ala Ser Lys Thr
 65                  70                  75                  80

Asn Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Thr Gln
                85                  90                  95

Ala Ile Val Arg Glu Gly Ile Lys Asn Val Thr Ala Gly Ala Asn Pro
            100                 105                 110

Ile Gly Ile Arg Arg Gly Ile Glu Thr Ala Val Ala Ala Ala Val Glu
        115                 120                 125

Ala Leu Lys Asn Asn Val Ile Pro Val Ala Asn Lys Glu Ala Ile Ala
    130                 135                 140

Gln Val Ala Ala Val Ser Ser Arg Ser Glu Lys Val Gly Glu Tyr Ile
145                 150                 155                 160

Ser Glu Ala Met Glu Lys Val Gly Lys Asp Gly Val Ile Thr Ile Glu
                165                 170                 175

Glu Ser Arg Gly Met Glu Thr Glu Leu Glu Val Val Glu Gly Met Gln
            180                 185                 190

Phe Asp Arg Gly Tyr Leu Ser Gln Tyr Met Val Thr Asp Ser Glu Lys
        195                 200                 205

Met Val Ala Asp Leu Glu Asn Pro Tyr Ile Leu Ile Thr Asp Lys Lys
    210                 215                 220

Ile Ser Asn Ile Gln Glu Ile Leu Pro Leu Leu Glu Ser Ile Leu Gln
225                 230                 235                 240

Ser Asn Arg Pro Leu Leu Ile Ile Ala Asp Asp Val Asp Gly Glu Ala
                245                 250                 255

Leu Pro Thr Leu Val Leu Asn Lys Ile Arg Gly Thr Phe Asn Val Val
            260                 265                 270

Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Glu
        275                 280                 285

Asp Ile Ala Ile Leu Thr Gly Gly Thr Val Ile Thr Glu Asp Leu Gly
    290                 295                 300

Leu Glu Leu Lys Asp Ala Thr Ile Glu Ala Leu Gly Gln Ala Ala Arg
305                 310                 315                 320

Val Thr Val Asp Lys Asp Ser Thr Val Ile Val Glu Gly Ala Gly Asn
                325                 330                 335

Pro Glu Ala Ile Ser His Arg Val Ala Val Ile Lys Ser Gln Ile Glu
            340                 345                 350
```

```
Thr Thr Thr Ser Glu Phe Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ser Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu
    370                 375                 380

Thr Glu Leu Lys Glu Met Lys Leu Arg Ile Glu Asp Ala Leu Asn Ala
385                 390                 395                 400

Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Thr Ala
                405                 410                 415

Leu Ala Asn Val Ile Pro Ala Val Ala Thr Leu Glu Leu Thr Gly Asp
            420                 425                 430

Glu Ala Thr Gly Arg Asn Ile Val Leu Arg Ala Leu Glu Glu Pro Val
        435                 440                 445

Arg Gln Ile Ala His Asn Ala Gly Phe Glu Gly Ser Ile Val Ile Asp
    450                 455                 460

Arg Leu Lys Asn Ala Glu Leu Gly Ile Gly Phe Asn Ala Ala Thr Gly
465                 470                 475                 480

Glu Trp Val Asn Met Ile Asp Gln Gly Ile Ile Asp Pro Val Lys Val
                485                 490                 495

Ser Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu Ile Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Ala Asn Lys Pro Glu Pro Val Ala Pro Ala
        515                 520                 525

Pro Ala Met Asp Pro Ser Met Met Gly Gly Met Gly Gly
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1662 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 15..1646

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
GAATTCGGCT TCAT ATG GCG GCT AAA GAT GTA AAA TTC GGT AAC GAC GCT
                Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala
```

50

```
                1               5                  10

CGT GTA AAA ATG CTC CGC GGC GTA AAC GTA CTG GCA GAC GCA GTT AAA          98
Arg Val Lys Met Leu Arg Gly Val Asn Val Leu Ala Asp Ala Val Lys
         15                  20                  25

GTA ACC CTG GGC CCG AAA GGC CGT AAC GTA GTG CTG GAC AAA TCC TTC         146
Val Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe
     30                  35                  40

GGC GCG CCA ACC ATC ACG AAA GAT GGT GTT TCT GTA GCA CGT GAA ATC         194
Gly Ala Pro Thr Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile
 45                  50                  55                  60

GAG CTG GAA GAC AAG TTC GAA AAC ATG GGC GCG CAG ATG GTG AAA GAA         242
Glu Leu Glu Asp Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu
                 65                  70                  75

GTG GCC TCT AAA GCG AAC GAC GCT GCA GGC GAC GGT ACC ACC ACC GCG         290
Val Ala Ser Lys Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala
             80                  85                  90

ACC GTG CTG GCT CAG GCT ATC ATC ACC GAA GGT CTG AAA GCC GTT GCT         338
Thr Val Leu Ala Gln Ala Ile Ile Thr Glu Gly Leu Lys Ala Val Ala
         95                  100                 105

GCG GGC ATG AAC CCA ATG GAT CTG AAA CGT GGT ATC GAC AAA GCT GTC         386
Ala Gly Met Asn Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val
     110                 115                 120

GCG TCC GCT GTT GAA GAA CTG AAA GCG CTG TCC GTA CCG TGC TCT GAC         434
Ala Ser Ala Val Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp
125                 130                 135                 140

TCT AAA GCC ATT GCT CAG GTA GGT ACC ATC TCC GCT AAC TCC GAC GAA         482
Ser Lys Ala Ile Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Glu
                145                 150                 155

ACC GTA GGT AAA CTG ATC GCG GAA GCG ATG GAT AAA GTC GGT AAA GAA         530
Thr Val Gly Lys Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys Glu
            160                 165                 170

GGC GTG ATC ACC GTT GAA GAC GGT ACC GGT CTG GAA GAC GAA CTG GAC         578
Gly Val Ile Thr Val Glu Asp Gly Thr Gly Leu Glu Asp Glu Leu Asp
        175                 180                 185

GTG GTT GAA GGT ATG CAG TTC GAC CGC GGT TAC CTG TCC CCA TAC TTC         626
Val Val Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe
    190                 195                 200

ATC AAC AAG CCA GAA ACT GGC GCT GTT GAG CTG GAA AGC CCG TTC ATC         674
Ile Asn Lys Pro Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile
205                 210                 215                 220

CTG CTG GCT GAC AAG AAA ATC TCC AAC ATC CGC GAA ATG CTG CCA GTG         722
```

```
Leu Leu Ala Asp Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val
                225                 230                 235

CTG GAA GCC GTT GCG AAA GCA GGC AAA CCG CTG GTT ATC ATT GCT GAA      770
Leu Glu Ala Val Ala Lys Ala Gly Lys Pro Leu Val Ile Ile Ala Glu
            240                 245                 250

GAC GTT GAA GGC GAA GCG CTG GCG ACC CTG GTG GTT AAC ACC ATG CGT      818
Asp Val Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg
        255                 260                 265

GGC ATC GTG AAA GTG GCT GCG GTT AAA GCA CCT GGC TTC GGC GAC CGC      866
Gly Ile Val Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg
    270                 275                 280

CGT AAA GCG ATG CTG CAG GAT ATC GCT ACC CTG ACC GGC GGT ACC GTC      914
Arg Lys Ala Met Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val
285                 290                 295                 300

ATC TCT GAA GAG ATC GGT ATG GAG CTG GAA AAA GCG ACC CTG GAA GAC      962
Ile Ser Glu Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp
                305                 310                 315

CTG GGC CAG GCT AAA CGT GTT GTG ATC AAC AAA GAC ACC ACC ACC ATC     1010
Leu Gly Gln Ala Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile
            320                 325                 330

ATC GAT GGC GTG GGC GAC GAA GCG GCG ATT CAG GGC CGT GTT GGT CAG     1058
Ile Asp Gly Val Gly Asp Glu Ala Ala Ile Gln Gly Arg Val Gly Gln
        335                 340                 345

ATC CGT AAG CAG ATC GAA GAA GCC ACT TCC GAT TAC GAC CGT GAA AAA     1106
Ile Arg Lys Gln Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys
    350                 355                 360

CTG CAG GAG CGC GTA GCG AAA CTG GCA GGC GGT GTT GCG GTA ATC AAA     1154
Leu Gln Glu Arg Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys
365                 370                 375                 380

GTC GGT GCT GCG ACT GAA GTT GAA ATG AAA GAG AAA AAA GCA CGC GTT     1202
Val Gly Ala Ala Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val
                385                 390                 395

GAC GAT GCC CTG CAC GCG ACC CGT GCT GCG GTA GAA GAA GGC GTG GTT     1250
Asp Asp Ala Leu His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val
            400                 405                 410

GCT GGT GGT GGT GTG GCG CTG GTG CGT GTT GCC GCG AAA CTG TCC GGC     1298
Ala Gly Gly Gly Val Ala Leu Val Arg Val Ala Ala Lys Leu Ser Gly
        415                 420                 425

CTG ACT GCT CAG AAC GAA GAT CAG AAC GTG GGT ATC AAA GTT GCG CTG     1346
Leu Thr Ala Gln Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu
    430                 435                 440
```

```
CGC GCA ATG GAA GCT CCA CTG CGT CAG ATC GTG TCC AAC GCC GGT GAA      1394
Arg Ala Met Glu Ala Pro Leu Arg Gln Ile Val Ser Asn Ala Gly Glu
445                 450                 455                 460

GAG CCA TCT GTT GTG ACC AAC AAC GTG AAA GCA GGC GAA GGT AAC TAC      1442
Glu Pro Ser Val Val Thr Asn Asn Val Lys Ala Gly Glu Gly Asn Tyr
                465                 470                 475

GGT TAC AAC GCA GCA ACT GAA GAA TAC GGC AAC ATG ATC GAC TTC GGT      1490
Gly Tyr Asn Ala Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Phe Gly
            480                 485                 490

ATC CTG GAT CCA ACC AAA GTG ACC CGT TCT GCT CTG CAG TAC GCG GCA      1538
Ile Leu Asp Pro Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala
        495                 500                 505

TCT GTC GCT GGC CTG ATG ATC ACC ACC GAG TGC ATG GTG ACC GAC CTG      1586
Ser Val Ala Gly Leu Met Ile Thr Thr Glu Cys Met Val Thr Asp Leu
    510                 515                 520

CCT AAA GGC GAC GCA CCT GAC TTA GGT GCT GCA GGC ATG GGT GGG ATG      1634
Pro Lys Gly Asp Ala Pro Asp Leu Gly Ala Ala Gly Met Gly Gly Met
525                 530                 535                 540

GGC GGT ATG ATG TGATCAAGCC GAATTC                                    1662
Gly Gly Met Met
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 544 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met
  1               5                  10                  15

Leu Arg Gly Val Asn Val Leu Ala Asp Ala Val Lys Val Thr Leu Gly
            20                  25                  30

Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala Pro Thr
        35                  40                  45

Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile Glu Leu Glu Asp
    50                  55                  60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys
65                  70                  75                  80
```

```
Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
                  85                  90                  95

Gln Ala Ile Ile Thr Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn
            100                 105                 110

Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Ala Ser Ala Val
        115                 120                 125

Glu Glu Leu Lys Ala Leu Ser Val Pro Cys Ser Asp Ser Lys Ala Ile
    130                 135                 140

Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Glu Thr Val Gly Lys
145                 150                 155                 160

Leu Ile Ala Glu Ala Met Asp Lys Val Gly Lys Glu Gly Val Ile Thr
                165                 170                 175

Val Glu Asp Gly Thr Gly Leu Glu Asp Glu Leu Asp Val Val Glu Gly
            180                 185                 190

Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Lys Pro
        195                 200                 205

Glu Thr Gly Ala Val Glu Leu Glu Ser Pro Phe Ile Leu Leu Ala Asp
    210                 215                 220

Lys Lys Ile Ser Asn Ile Arg Glu Met Leu Pro Val Leu Glu Ala Val
225                 230                 235                 240

Ala Lys Ala Gly Lys Pro Leu Val Ile Ile Ala Glu Asp Val Glu Gly
                245                 250                 255

Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile Val Lys
            260                 265                 270

Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met
        275                 280                 285

Leu Gln Asp Ile Ala Thr Leu Thr Gly Gly Thr Val Ile Ser Glu Glu
    290                 295                 300

Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Asp Leu Gly Gln Ala
305                 310                 315                 320

Lys Arg Val Val Ile Asn Lys Asp Thr Thr Thr Ile Ile Asp Gly Val
                325                 330                 335

Gly Asp Glu Ala Ala Ile Gln Gly Arg Val Gly Gln Ile Arg Lys Gln
            340                 345                 350

Ile Glu Glu Ala Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg
        355                 360                 365

Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala
```

```
    370                 375                 380

Thr Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val Asp Asp Ala Leu
385                 390                 395                 400

His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly
                405                 410                 415

Val Ala Leu Val Arg Val Ala Ala Lys Leu Ser Gly Leu Thr Ala Gln
            420                 425                 430

Asn Glu Asp Gln Asn Val Gly Ile Lys Val Ala Leu Arg Ala Met Glu
        435                 440                 445

Ala Pro Leu Arg Gln Ile Val Ser Asn Ala Gly Glu Glu Pro Ser Val
    450                 455                 460

Val Thr Asn Asn Val Lys Ala Gly Glu Gly Asn Tyr Gly Tyr Asn Ala
465                 470                 475                 480

Ala Thr Glu Glu Tyr Gly Asn Met Ile Asp Phe Gly Ile Leu Asp Pro
                485                 490                 495

Thr Lys Val Thr Arg Ser Ala Leu Gln Tyr Ala Ala Ser Val Ala Gly
            500                 505                 510

Leu Met Ile Thr Thr Glu Cys Met Val Thr Asp Leu Pro Lys Gly Asp
        515                 520                 525

Ala Pro Asp Leu Gly Ala Ala Gly Met Gly Gly Met Gly Gly Met Met
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1661 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 15..1649

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
GAATTCGGCT TCAT ATG GCA AAA GAA ATC AAA TTT TCA GCA GAT GCG CGT        50
                Met Ala Lys Glu Ile Lys Phe Ser Ala Asp Ala Arg
                  1               5                  10

GCT GCC ATG GTG CGC GGA GTT GAT ATG TTA GCA GAT ACC GTC AAA GTA        98
Ala Ala Met Val Arg Gly Val Asp Met Leu Ala Asp Thr Val Lys Val
```

```
        15                  20                  25

ACG CTT GGT CCT AAA GGG CGC AAT GTT GTT CTT GAA AAA GCT TTT GGT      146
Thr Leu Gly Pro Lys Gly Arg Asn Val Val Leu Glu Lys Ala Phe Gly
     30                  35                  40

TCT CCC TTA ATT ACT AAT GAC GGG GTA ACC ATT GCT AAA GAG ATC GAA      194
Ser Pro Leu Ile Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu
 45                  50                  55                  60

TTA GAA GAT CAT TTT GAA AAC ATG GGA GCA AAA TTG GTG TCT GAA GTG      242
Leu Glu Asp His Phe Glu Asn Met Gly Ala Lys Leu Val Ser Glu Val
                 65                  70                  75

GCT TCT AAA ACC AAT GAT ATT GCT GGT GAT GGG ACG ACT ACT GCA ACA      290
Ala Ser Lys Thr Asn Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr
             80                  85                  90

GTT TTG ACA CAA GCC ATT GTT CAT GAA GGA CTA AAA AAT GTG ACA GCA      338
Val Leu Thr Gln Ala Ile Val His Glu Gly Leu Lys Asn Val Thr Ala
         95                  100                 105

GGT GCT AAT CCA ATT GGT ATC CGT CGA GGC ATT GAA ACA GCA ACA GCA      386
Gly Ala Asn Pro Ile Gly Ile Arg Arg Gly Ile Glu Thr Ala Thr Ala
     110                 115                 120

ACA GCT GTT GAA GCC TTG AAA GCC ATT GCT CAA CCT GTA TCT GGC AAG      434
Thr Ala Val Glu Ala Leu Lys Ala Ile Ala Gln Pro Val Ser Gly Lys
125                 130                 135                 140

GAA GCT ATT GCT CAG GTC GCT GCA GTA TCA TCA CGC TCT GAA AAA GTT      482
Glu Ala Ile Ala Gln Val Ala Ala Val Ser Ser Arg Ser Glu Lys Val
                145                 150                 155

GGA GAG TAT ATC TCA GAA GCT ATG GAG CGT GTG GGC AAC GAT GGT GTG      530
Gly Glu Tyr Ile Ser Glu Ala Met Glu Arg Val Gly Asn Asp Gly Val
            160                 165                 170

ATT ACC ATC GAA GAA TCT CGA GGT ATG GAA ACA GAA CTT GAA GTG GTT      578
Ile Thr Ile Glu Glu Ser Arg Gly Met Glu Thr Glu Leu Glu Val Val
        175                 180                 185

GAA GGC ATG CAA TTT GAC CGT GGT TAC CTG TCT CAA TAC ATG GTC ACA      626
Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Gln Tyr Met Val Thr
    190                 195                 200

GAC AAT GAA AAA ATG GTT GCA GAC CTT GAA AAC CCA TTT ATC TTA ATC      674
Asp Asn Glu Lys Met Val Ala Asp Leu Glu Asn Pro Phe Ile Leu Ile
205                 210                 215                 220

ACG GAT AAA AAA GTG TCA AAC ATC CAA GAC ATT TTG CCA CTA CTT GAG      722
Thr Asp Lys Lys Val Ser Asn Ile Gln Asp Ile Leu Pro Leu Leu Glu
                225                 230                 235

GAA GTT CTT AAA ACC AAC CGT CCA TTA CTC ATT ATT GCA GAT GAT GTG      770
```

```
Glu Val Leu Lys Thr Asn Arg Pro Leu Leu Ile Ile Ala Asp Asp Val
            240                 245                 250

GAT GGT GAA GCA CTT CCA ACC CTT GTC TTG AAC AAG ATT CGT GGT ACT      818
Asp Gly Glu Ala Leu Pro Thr Leu Val Leu Asn Lys Ile Arg Gly Thr
        255                 260                 265

TTC AAT GTG GTT GCT GTC AAA GCG CCA GGA TTT GGT GAT CGT CGT AAA      866
Phe Asn Val Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys
    270                 275                 280

GCT ATG CTT GAA GAC ATT GCT ATC TTG ACA GGT GGT ACA GTG ATT ACA      914
Ala Met Leu Glu Asp Ile Ala Ile Leu Thr Gly Gly Thr Val Ile Thr
285                 290                 295                 300

GAG GAT CTA GGA CTT GAA TTA AAA GAT GCT ACA ATG ACA GCC CTT GGA      962
Glu Asp Leu Gly Leu Glu Leu Lys Asp Ala Thr Met Thr Ala Leu Gly
                305                 310                 315

CAG GCT GCT AAG ATT ACA GTT GAT AAA GAT AGC ACA GTA ATT GTT GAA     1010
Gln Ala Ala Lys Ile Thr Val Asp Lys Asp Ser Thr Val Ile Val Glu
            320                 325                 330

GGT TCA GGA AGT TCA GAA GCT ATT GCT AAC CGT ATT GCA CTG ATT AAA     1058
Gly Ser Gly Ser Ser Glu Ala Ile Ala Asn Arg Ile Ala Leu Ile Lys
        335                 340                 345

TCG CAA TTA GAA ACA ACA ACT TCT GAC TTT GAC CGT GAA AAA CTA CAA     1106
Ser Gln Leu Glu Thr Thr Thr Ser Asp Phe Asp Arg Glu Lys Leu Gln
    350                 355                 360

GAA CGT TTG GCG AAA TTA GCT GGT GGT GTA GCT GTT ATC AAA GTA GGA     1154
Glu Arg Leu Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly
365                 370                 375                 380

GCT CCA ACA GAG ACA GCT TTA AAA GAA ATG AAA CTT CGC ATT GAG GAT     1202
Ala Pro Thr Glu Thr Ala Leu Lys Glu Met Lys Leu Arg Ile Glu Asp
                385                 390                 395

GCT CTA AAT GCT ACA CGT GCA GCC GTT GAA GAA GGT ATC GTT GCT GGT     1250
Ala Leu Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly
            400                 405                 410

GGT GGA ACA GCA CTT ATT ACG GTT ATT GAA AAA GTA GCA GCT CTT GAG     1298
Gly Gly Thr Ala Leu Ile Thr Val Ile Glu Lys Val Ala Ala Leu Glu
        415                 420                 425

CTT GAG GGC GAT GAT GCT ACT GGA CGT AAC ATT GTG CTT CGT GCT CTA     1346
Leu Glu Gly Asp Asp Ala Thr Gly Arg Asn Ile Val Leu Arg Ala Leu
    430                 435                 440

GAA GAG CCT GTA CGT CAA ATT GCT TTA AAT GCT GGG TAC GAA GGC TCC     1394
Glu Glu Pro Val Arg Gln Ile Ala Leu Asn Ala Gly Tyr Glu Gly Ser
445                 450                 455                 460
```

```
GTA GTT ATT GAC AAG TTG AAA AAC AGC CCT GCA GGA ACA GGA TTT AAT      1442
Val Val Ile Asp Lys Leu Lys Asn Ser Pro Ala Gly Thr Gly Phe Asn
                465                 470                 475

GCT GCA ACA GGT GAG TGG GTT GAT ATG ATT AAA ACA GGA ATC ATT GAC      1490
Ala Ala Thr Gly Glu Trp Val Asp Met Ile Lys Thr Gly Ile Ile Asp
            480                 485                 490

CCT GTC AAA GTA ACA CGA TCA GCG CTT CAA AAT GCA GCT TCT GTA GCT      1538
Pro Val Lys Val Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala
        495                 500                 505

AGT CTT ATT TTG ACA ACA GAA GCA GTT GTT GCT AAT AAA CCT GAA CCA      1586
Ser Leu Ile Leu Thr Thr Glu Ala Val Val Ala Asn Lys Pro Glu Pro
    510                 515                 520

GCT ACG CCA GCG CCA GCA ATG CCA GCA GGT ATG GAT CCA GGA ATG ATG      1634
Ala Thr Pro Ala Pro Ala Met Pro Ala Gly Met Asp Pro Gly Met Met
525                 530                 535                 540

GGT GGG ATG GGC GGA TAAGCCGAAT TC                                   1661
Gly Gly Met Gly Gly
                545
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 545 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Ala Lys Glu Ile Lys Phe Ser Ala Asp Ala Arg Ala Ala Met Val
  1               5                  10                  15

Arg Gly Val Asp Met Leu Ala Asp Thr Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Ala Phe Gly Ser Pro Leu Ile
        35                  40                  45

Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp His
    50                  55                  60

Phe Glu Asn Met Gly Ala Lys Leu Val Ser Glu Val Ala Ser Lys Thr
65                  70                  75                  80

Asn Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Thr Gln
                85                  90                  95

Ala Ile Val His Glu Gly Leu Lys Asn Val Thr Ala Gly Ala Asn Pro
```

```
            100                 105                 110

Ile Gly Ile Arg Arg Gly Ile Glu Thr Ala Thr Ala Thr Ala Val Glu
        115                 120                 125

Ala Leu Lys Ala Ile Ala Gln Pro Val Ser Gly Lys Glu Ala Ile Ala
    130                 135                 140

Gln Val Ala Ala Val Ser Ser Arg Ser Glu Lys Val Gly Glu Tyr Ile
145                 150                 155                 160

Ser Glu Ala Met Glu Arg Val Gly Asn Asp Gly Val Ile Thr Ile Glu
                165                 170                 175

Glu Ser Arg Gly Met Glu Thr Glu Leu Glu Val Val Glu Gly Met Gln
            180                 185                 190

Phe Asp Arg Gly Tyr Leu Ser Gln Tyr Met Val Thr Asp Asn Glu Lys
        195                 200                 205

Met Val Ala Asp Leu Glu Asn Pro Phe Ile Leu Ile Thr Asp Lys Lys
    210                 215                 220

Val Ser Asn Ile Gln Asp Ile Leu Pro Leu Leu Glu Glu Val Leu Lys
225                 230                 235                 240

Thr Asn Arg Pro Leu Leu Ile Ile Ala Asp Asp Val Asp Gly Glu Ala
                245                 250                 255

Leu Pro Thr Leu Val Leu Asn Lys Ile Arg Gly Thr Phe Asn Val Val
            260                 265                 270

Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Glu
        275                 280                 285

Asp Ile Ala Ile Leu Thr Gly Gly Thr Val Ile Thr Glu Asp Leu Gly
    290                 295                 300

Leu Glu Leu Lys Asp Ala Thr Met Thr Ala Leu Gly Gln Ala Ala Lys
305                 310                 315                 320

Ile Thr Val Asp Lys Asp Ser Thr Val Ile Val Glu Gly Ser Gly Ser
                325                 330                 335

Ser Glu Ala Ile Ala Asn Arg Ile Ala Leu Ile Lys Ser Gln Leu Glu
            340                 345                 350

Thr Thr Thr Ser Asp Phe Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Pro Thr Glu
    370                 375                 380

Thr Ala Leu Lys Glu Met Lys Leu Arg Ile Glu Asp Ala Leu Asn Ala
385                 390                 395                 400
```

```
Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Thr Ala
                405                 410                 415

Leu Ile Thr Val Ile Glu Lys Val Ala Ala Leu Glu Leu Glu Gly Asp
            420                 425                 430

Asp Ala Thr Gly Arg Asn Ile Val Leu Arg Ala Leu Glu Glu Pro Val
        435                 440                 445

Arg Gln Ile Ala Leu Asn Ala Gly Tyr Glu Gly Ser Val Val Ile Asp
    450                 455                 460

Lys Leu Lys Asn Ser Pro Ala Gly Thr Gly Phe Asn Ala Ala Thr Gly
465                 470                 475                 480

Glu Trp Val Asp Met Ile Lys Thr Gly Ile Ile Asp Pro Val Lys Val
                485                 490                 495

Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu Ile Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Ala Asn Lys Pro Glu Pro Ala Thr Pro Ala
        515                 520                 525

Pro Ala Met Pro Ala Gly Met Asp Pro Gly Met Met Gly Gly Met Gly
    530                 535                 540

Gly
545
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 544 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Ala Lys Glu Ile Lys Phe Ser Glu Glu Ala Arg Arg Ala Met Leu
1               5                   10                  15

Arg Gly Val Asp Ala Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Lys Phe Gly Ser Pro Leu Ile
        35                  40                  45

Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp Ala
```

```
    50                  55                  60

Phe Glu Asn Met Gly Ala Lys Leu Val Ala Glu Val Ala Ser Lys Thr
65                  70                  75                  80

Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
                85                  90                  95

Ala Met Ile Arg Glu Gly Leu Lys Asn Val Thr Ala Gly Ala Asn Pro
            100                 105                 110

Val Gly Val Arg Lys Gly Met Glu Gln Ala Val Ala Val Ala Ile Glu
        115                 120                 125

Asn Leu Lys Glu Ile Ser Lys Pro Ile Glu Gly Lys Glu Ser Ile Ala
    130                 135                 140

Gln Val Ala Ala Ile Ser Ala Ala Asp Glu Glu Val Gly Ser Leu Ile
145                 150                 155                 160

Ala Glu Ala Met Glu Arg Val Gly Asn Asp Gly Val Ile Thr Ile Glu
                165                 170                 175

Glu Ser Lys Gly Phe Thr Thr Glu Leu Glu Val Val Glu Gly Met Gln
            180                 185                 190

Phe Asp Arg Gly Tyr Ala Ser Pro Tyr Met Val Thr Asp Ser Asp Lys
        195                 200                 205

Met Glu Ala Val Leu Asp Asn Pro Tyr Ile Leu Ile Thr Asp Lys Lys
    210                 215                 220

Ile Thr Asn Ile Gln Glu Ile Leu Pro Val Leu Glu Gln Val Val Gln
225                 230                 235                 240

Gln Gly Lys Pro Leu Leu Leu Ile Ala Glu Asp Val Glu Gly Glu Ala
                245                 250                 255

Leu Ala Thr Leu Val Val Asn Lys Leu Arg Gly Thr Phe Asn Ala Val
            260                 265                 270

Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Glu
        275                 280                 285

Asp Ile Ala Val Leu Thr Gly Gly Glu Val Ile Thr Glu Asp Leu Gly
    290                 295                 300

Leu Asp Leu Lys Ser Thr Gln Ile Ala Gln Leu Gly Arg Ala Ser Lys
305                 310                 315                 320

Val Val Val Thr Lys Glu Asn Thr Thr Ile Val Glu Gly Ala Gly Glu
                325                 330                 335

Thr Asp Lys Ile Ser Ala Arg Val Thr Gln Ile Arg Ala Gln Val Glu
            340                 345                 350
```

```
Glu Thr Thr Ser Glu Phe Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu
    370                 375                 380

Thr Glu Leu Lys Glu Arg Lys Leu Arg Ile Glu Asp Ala Leu Asn Ser
385                 390                 395                 400

Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ser Gly Gly Gly Thr Ala
                405                 410                 415

Leu Val Asn Val Tyr Asn Lys Val Ala Ala Val Glu Ala Glu Gly Asp
            420                 425                 430

Ala Gln Thr Gly Ile Asn Ile Val Leu Arg Ala Leu Glu Glu Pro Ile
        435                 440                 445

Arg Gln Ile Ala His Asn Ala Gly Leu Glu Gly Ser Val Ile Val Glu
    450                 455                 460

Arg Leu Lys Asn Glu Glu Ile Gly Val Gly Phe Asn Ala Ala Thr Gly
465                 470                 475                 480

Glu Trp Val Asn Met Ile Glu Lys Gly Ile Val Asp Pro Thr Lys Val
                485                 490                 495

Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ala Met Phe Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Ala Asp Lys Pro Glu Glu Asn Gly Gly Gly
        515                 520                 525

Ala Gly Met Pro Asp Met Gly Gly Met Gly Gly Met Gly Gly Met Met
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 539 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Met Ala Lys Thr Leu Leu Phe Gly Glu Glu Ala Arg Arg Ser Met Gln
1               5                   10                  15
```

```
Ala Gly Val Asp Lys Leu Ala Asn Thr Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Ile Leu Asp Lys Lys Phe Gly Ser Pro Leu Ile
        35                  40                  45

Thr Asn Asp Gly Val Thr Ile Ala Arg Glu Ile Glu Leu Glu Asp Ala
    50                  55                  60

Tyr Glu Asn Met Gly Ala Gln Leu Val Lys Glu Val Ala Thr Lys Thr
65                  70                  75                  80

Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Leu Leu Ala Gln
                85                  90                  95

Ala Ile Ile Arg Glu Gly Leu Lys Asn Val Thr Ala Gly Ala Asn Pro
            100                 105                 110

Ile Leu Ile Arg Asn Gly Ile Lys Thr Ala Val Glu Lys Ala Val Glu
        115                 120                 125

Glu Ile Gln Lys Ile Ser Lys Pro Val Asn Gly Lys Glu Asp Ile Ala
    130                 135                 140

Arg Val Ala Ala Ile Ser Ala Ala Asp Glu Lys Ile Gly Lys Leu Ile
145                 150                 155                 160

Ala Asp Ala Met Glu Lys Val Gly Asn Glu Gly Val Ile Thr Val Glu
                165                 170                 175

Glu Ser Lys Ser Met Gly Thr Glu Leu Asp Val Val Glu Gly Met Gln
            180                 185                 190

Phe Asp Arg Gly Tyr Val Ser Ala Tyr Met Val Thr Asp Thr Glu Lys
        195                 200                 205

Met Glu Ala Val Leu Asp Asn Pro Leu Val Leu Ile Thr Asp Lys Lys
    210                 215                 220

Ile Ser Asn Ile Gln Asp Leu Leu Pro Leu Leu Glu Gln Ile Val Gln
225                 230                 235                 240

Ala Gly Lys Lys Leu Leu Ile Ile Ala Asp Asp Ile Glu Gly Glu Ala
                245                 250                 255

Met Thr Thr Leu Val Val Asn Lys Leu Arg Gly Thr Phe Thr Cys Val
            260                 265                 270

Gly Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Glu Met Leu Gln
        275                 280                 285

Asp Ile Ala Thr Leu Thr Gly Gly Val Val Ile Ser Asp Glu Val Gly
    290                 295                 300

Gly Asp Leu Lys Glu Ala Thr Leu Asp Met Leu Gly Glu Ala Glu Ser
```

```
305                 310                 315                 320

Val Lys Val Thr Lys Glu Ser Thr Thr Ile Val Asn Gly Arg Gly Asn
                325                 330                 335

Ser Glu Glu Ile Lys Asn Arg Ile Asn Gln Ile Lys Leu Gln Leu Glu
            340                 345                 350

Ala Thr Thr Ser Glu Phe Asp Lys Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Val Lys Val Gly Ala Ala Thr Glu
    370                 375                 380

Thr Glu Leu Lys Glu Ser Lys Leu Arg Ile Glu Asp Ala Leu Ala Ala
385                 390                 395                 400

Thr Lys Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly Thr Ala
                405                 410                 415

Tyr Val Asn Val Ile Asn Glu Val Ala Lys Leu Thr Ser Asp Ile Gln
            420                 425                 430

Asp Glu Gln Val Gly Ile Asn Ile Ile Val Arg Ser Leu Glu Glu Pro
        435                 440                 445

Met Arg Gln Ile Ala His Asn Ala Gly Leu Glu Gly Ser Val Ile Ile
    450                 455                 460

Glu Lys Val Lys Asn Ser Asp Ala Gly Val Gly Phe Asp Ala Leu Arg
465                 470                 475                 480

Gly Glu Tyr Lys Asp Met Ile Lys Ala Gly Ile Val Asp Pro Thr Lys
                485                 490                 495

Val Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Thr Phe
            500                 505                 510

Leu Thr Thr Glu Ala Ala Val Ala Asp Ile Pro Glu Lys Glu Met Pro
        515                 520                 525

Gln Gly Ala Gly Met Gly Met Asp Gly Met Tyr
    530                 535
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 551 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Met Ala Asn Met Val Val Thr Gly Glu Gln Leu Asp Lys Ser Ile Arg
1               5                  10                  15

Glu Val Val Arg Ile Leu Glu Asp Ala Val Gly Cys Thr Ala Gly Pro
            20                  25                  30

Lys Gly Leu Thr Val Ala Ile Ser Lys Pro Tyr Gly Ala Pro Glu Val
        35                  40                  45

Thr Lys Asp Gly Tyr Lys Val Met Lys Ser Ile Lys Pro Glu Asp Pro
    50                  55                  60

Leu Ala Leu Ala Ile Ala Asn Ile Ile Ala Gln Ser Ala Ser Gln Cys
65                  70                  75                  80

Asn Asp Lys Val Gly Asp Gly Thr Thr Thr Cys Ser Ile Leu Thr Ala
                85                  90                  95

Lys Val Ile Glu Glu Val Ser Lys Val Lys Ala Ala Gly Ala Asp Ile
            100                 105                 110

Ile Cys Val Arg Glu Gly Val Leu Lys Ala Lys Glu Ala Val Leu Glu
        115                 120                 125

Ala Leu Lys Cys Met Lys Arg Glu Val Leu Ser Glu Glu Glu Ile Ala
    130                 135                 140

Gln Val Ala Thr Ile Ser Ala Asn Gly Asp Lys Asn Ile Gly Thr Lys
145                 150                 155                 160

Ile Ala Gln Cys Val Lys Glu Val Gly Lys Asp Gly Val Ile Thr Val
                165                 170                 175

Glu Glu Ser Lys Gly Phe Lys Glu Leu Asp Val Glu Lys Thr Asp Gly
            180                 185                 190

Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Val Thr Asn Ser
        195                 200                 205

Glu Lys Met Leu Val Glu Phe Glu Asn Pro Tyr Ile Leu Leu Thr Glu
    210                 215                 220

Lys Lys Leu Asn Ile Ile Gln Pro Leu Leu Pro Ile Leu Glu Asn Ile
225                 230                 235                 240

Ala Arg Ser Gly Arg Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly
                245                 250                 255

Glu Ala Leu Ser Thr Leu Val Leu Asn Lys Leu Arg Gly Gly Leu His
            260                 265                 270

Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Asp Met
```

```
        275                 280                 285

Leu Gly Asp Ile Ala Ile Leu Thr Gly Ala Lys His Val Ile Asn Asp
    290                 295                 300

Glu Leu Ala Ile Lys Met Glu Asp Leu Thr Leu Cys Asp Leu Gly Thr
305                 310                 315                 320

Ala Lys Asn Ile Arg Ile Thr Lys Asp Thr Thr Thr Ile Ile Gly Ser
                325                 330                 335

Val Asp Asn Ser Cys Ala His Val Gln Ser Arg Ile Cys Gln Ile Arg
            340                 345                 350

Met Gln Ile Asp Asn Ser Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln
        355                 360                 365

Glu Arg Leu Ala Lys Leu Ser Gly Gly Val Ala Val Leu Lys Val Gly
    370                 375                 380

Gly Ser Ser Glu Val Glu Val Lys Glu Arg Lys Asp Arg Val Glu Asp
385                 390                 395                 400

Ala Leu His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Pro Gly
                405                 410                 415

Gly Gly Ala Ala Leu Leu Tyr Thr Leu Ser Ala Leu Asp Asn Leu Lys
            420                 425                 430

Ser Lys Asn Asp Asp Glu Gln Leu Gly Ile Asn Ile Val Lys Arg Ala
        435                 440                 445

Leu Gln Ala Pro Ile Lys Arg Ile Ile Lys Asn Ala Gly Ser Glu Asn
    450                 455                 460

Ala Pro Cys Val Ile Ala His Leu Leu Lys Gln Asn Asp Lys Glu Leu
465                 470                 475                 480

Ile Phe Asn Val Asp Val Thr Asn Phe Ala Asn Ala Phe Thr Ser Gly
                485                 490                 495

Val Ile Asp Pro Leu Lys Val Val Arg Ile Ala Phe Asp Phe Ala Val
            500                 505                 510

Ser Leu Ala Ala Val Phe Met Thr Leu Asn Ala Ile Val Val Asp Ile
        515                 520                 525

Pro Ser Lys Asp Asp Asn Ser Ala Ala Gly Gly Ala Gly Met Gly Gly
    530                 535                 540

Met Gly Gly Met Gly Gly Phe
545                 550
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 548 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Ala Ala Lys Asp Val Lys Phe Gly Asn Asp Ala Arg Val Lys Met
1               5                  10                  15

Leu Asn Gly Val Asn Ile Leu Ala Asp Ala Val Lys Val Thr Leu Gly
            20                  25                  30

Pro Lys Gly Arg Asn Val Val Leu Asp Lys Ser Phe Gly Ala Pro Thr
        35                  40                  45

Ile Thr Lys Asp Gly Val Ser Val Ala Arg Glu Ile Glu Leu Glu Asp
    50                  55                  60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys
65                  70                  75                  80

Ala Asn Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
                85                  90                  95

Gln Ala Ile Val Asn Glu Gly Leu Lys Ala Val Ala Ala Gly Met Asn
            100                 105                 110

Pro Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Asn Ser Val Val
        115                 120                 125

Ala Glu Leu Lys Asn Leu Ser Lys Pro Cys Glu Thr Ser Lys Glu Ile
    130                 135                 140

Glu Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Ser Ile Val Gly Gln
145                 150                 155                 160

Leu Ile Ala Gln Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr
                165                 170                 175

Val Glu Asp Gly Thr Gly Leu Glu Asp Glu Leu Asp Val Val Glu Gly
            180                 185                 190

Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Lys Pro
        195                 200                 205

Glu Thr Ala Gly Thr Val Glu Leu Asp Asn Pro Phe Ile Leu Leu Val
    210                 215                 220

Asp Lys Lys Ile Ser Asn Ile Arg Glu Leu Leu Pro Val Leu Glu Ala
```

225 230 235 240

Val Ala Lys Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val Glu
245 250 255

Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr Met Arg Gly Ile Val
260 265 270

Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala
275 280 285

Met Leu Gln Asp Ile Ala Ile Leu Thr Ala Gly Thr Val Ile Ser Glu
290 295 300

Glu Ile Gly Met Glu Leu Glu Lys Ala Thr Leu Glu Glu Leu Gly Gln
305 310 315 320

Ala Lys Arg Val Val Ile Thr Lys Asp Asn Thr Thr Ile Ile Asp Gly
325 330 335

Ile Gly Asp Glu Ala Gln Ile Lys Ala Arg Val Val Gln Ile Arg Gln
340 345 350

Gln Ile Glu Asp Ser Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu
355 360 365

Arg Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala
370 375 380

Ala Thr Glu Val Ala Met Lys Glu Lys Lys Asp Arg Val Asp Asp Ala
385 390 395 400

Leu His Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly
405 410 415

Gly Val Ala Leu Val Arg Ala Ala Asn Lys Val Ser Ala Thr Leu Thr
420 425 430

Gly Asp Asn Glu Glu Gln Asn Val Gly Ile Lys Leu Ala Leu Arg Ala
435 440 445

Met Glu Ala Pro Leu Arg Gln Ile Val Glu Asn Ser Gly Glu Asp Ala
450 455 460

Ser Val Val Ala Arg Asp Val Lys Asp Gly Ser Gly Asn Phe Gly Tyr
465 470 475 480

Asn Ala Thr Thr Glu Glu Tyr Gly Asp Met Leu Glu Met Gly Ile Leu
485 490 495

Asp Pro Thr Lys Val Thr Arg Ser Ala Leu Gln Phe Ala Ala Ser Ile
500 505 510

Ala Gly Leu Met Ile Thr Thr Glu Cys Met Ile Thr Asp Leu Pro Lys
515 520 525

```
Glu Asp Lys Leu Asp Ala Gln Ala Ala Met Gly Gly Met Gly Gly Met
    530                 535                 540

Gly Gly Met Met
545
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 549 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Ala Lys Glu Leu Arg Phe Gly Asp Asp Ala Arg Leu Gln Met Leu
1               5                   10                  15

Ala Gly Val Asn Ala Leu Ala Asp Ala Val Gln Val Thr Met Gly Pro
            20                  25                  30

Arg Gly Arg Asn Val Val Leu Glu Lys Ser Tyr Gly Ala Pro Thr Val
        35                  40                  45

Thr Lys Asp Gly Val Ser Val Ala Lys Glu Ile Glu Phe Glu His Arg
    50                  55                  60

Phe Met Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys Thr
65                  70                  75                  80

Ser Asp Thr Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Arg
                85                  90                  95

Ser Ile Leu Val Glu Gly His Lys Ala Val Ala Ala Gly Met Asn Pro
            100                 105                 110

Met Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Leu Ala Val Thr Lys
        115                 120                 125

Lys Leu Gln Ala Met Ser Lys Pro Cys Lys Asp Ser Lys Ala Ile Ala
    130                 135                 140

Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Glu Ala Ile Gly Ala Ile
145                 150                 155                 160

Ile Ala Glu Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr Val
                165                 170                 175

Glu Asp Gly Asn Gly Leu Glu Asn Glu Leu Ser Val Val Glu Gly Met
```

```
            180                 185                 190

Gln Phe Asp Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Asn Asn Gln Gln
        195                 200                 205

Asn Met Ser Cys Glu Leu Glu His Pro Phe Ile Leu Leu Val Asp Lys
    210                 215                 220

Lys Val Ser Ser Ile Arg Glu Met Leu Ser Val Leu Glu Gly Val Ala
225                 230                 235                 240

Lys Ser Gly Arg Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly Glu
                245                 250                 255

Ala Leu Ala Thr Leu Val Val Asn Asn Met Arg Gly Ile Val Lys Val
            260                 265                 270

Cys Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu
        275                 280                 285

Gln Asp Ile Ala Ile Leu Thr Lys Gly Gln Val Ile Ser Glu Glu Ile
    290                 295                 300

Gly Lys Ser Leu Glu Gly Ala Thr Leu Glu Asp Leu Gly Ser Ala Lys
305                 310                 315                 320

Arg Ile Val Val Thr Lys Glu Asn Thr Thr Ile Ile Asp Gly Glu Gly
                325                 330                 335

Lys Ala Thr Glu Ile Asn Ala Arg Ile Thr Gln Ile Arg Ala Gln Met
            340                 345                 350

Glu Glu Thr Thr Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Val
        355                 360                 365

Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr
    370                 375                 380

Glu Val Glu Met Lys Glu Lys Lys Ala Arg Val Glu Asp Ala Leu His
385                 390                 395                 400

Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Val
                405                 410                 415

Ala Leu Ile Arg Ala Gln Lys Ala Leu Asp Ser Leu Lys Gly Asp Asn
            420                 425                 430

Asp Asp Gln Asn Met Gly Ile Asn Ile Leu Arg Arg Ala Ile Glu Ser
        435                 440                 445

Pro Met Arg Gln Ile Val Thr Asn Ala Gly Tyr Glu Ala Ser Val Val
    450                 455                 460

Val Asn Lys Val Ala Glu His Lys Asp Asn Tyr Gly Phe Asn Ala Ala
465                 470                 475                 480
```

Thr Gly Glu Tyr Gly Asp Met Val Glu Met Gly Ile Leu Asp Pro Thr
485 490 495

Lys Val Thr Arg Met Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu
500 505 510

Met Leu Thr Thr Glu Cys Met Val Ala Asp Leu Pro Lys Lys Glu Glu
515 520 525

Gly Val Gly Ala Gly Asp Met Gly Gly Met Gly Gly Met Gly Gly Met
530 535 540

Gly Gly Met Met Glx
545

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 541 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Lys Thr Ile Ala Tyr Asp Glu Glu Ala Arg Arg Gly Leu Glu
1 5 10 15

Arg Gly Leu Asn Ala Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro
20 25 30

Lys Gly Arg Asn Val Val Leu Glu Lys Lys Trp Gly Ala Pro Thr Ile
35 40 45

Thr Asn Asp Gly Val Ser Ile Ala Lys Glu Ile Glu Leu Glu Asp Pro
50 55 60

Tyr Glu Lys Ile Gly Ala Glu Leu Val Lys Glu Val Ala Lys Lys Thr
65 70 75 80

Asp Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
85 90 95

Ala Leu Val Arg Glu Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro
100 105 110

Leu Gly Leu Lys Arg Gly Ile Glu Lys Ala Val Glu Lys Val Thr Glu
115 120 125

Thr Leu Leu Lys Ser Ala Lys Glu Val Glu Thr Lys Asp Gln Ile Ala

```
    130                 135                 140

Ala Thr Ala Ala Ile Ser Ala Gly Asp Gln Ser Ile Gly Asp Leu Ile
145                 150                 155                 160

Ala Glu Ala Met Asp Lys Val Gly Asn Glu Gly Val Ile Thr Val Glu
                165                 170                 175

Glu Ser Asn Thr Phe Gly Leu Gln Leu Glu Leu Thr Glu Gly Met Arg
            180                 185                 190

Phe Asp Lys Gly Tyr Ile Ser Gly Tyr Phe Val Thr Asp Ala Glu Arg
        195                 200                 205

Gln Glu Ala Val Leu Glu Asp Pro Phe Ile Leu Leu Val Ser Ser Lys
    210                 215                 220

Val Ser Thr Val Lys Asp Leu Leu Pro Leu Leu Glu Lys Val Ile Gln
225                 230                 235                 240

Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly Glu Ala
                245                 250                 255

Leu Ser Thr Leu Val Val Asn Lys Ile Arg Gly Thr Phe Lys Ser Val
            260                 265                 270

Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Gln
        275                 280                 285

Asp Met Ala Ile Leu Thr Gly Gly Gln Val Ile Ser Glu Glu Val Gly
    290                 295                 300

Leu Ser Leu Glu Ser Ala Asp Ile Ser Leu Leu Gly Lys Ala Arg Lys
305                 310                 315                 320

Val Val Val Thr Lys Asp Glu Thr Thr Ile Val Glu Gly Ala Gly Asp
                325                 330                 335

Ser Asp Ala Ile Ala Gly Arg Val Ala Gln Ile Arg Thr Glu Ile Glu
            340                 345                 350

Asn Ser Asp Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Ala Gly Ala Ala Thr Glu
    370                 375                 380

Val Glu Leu Lys Glu Arg Lys His Arg Ile Glu Asp Ala Val Arg Asn
385                 390                 395                 400

Ala Lys Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Val Ala
                405                 410                 415

Leu Leu His Ala Ile Pro Ala Leu Asp Glu Leu Lys Pro Glu Gly Glu
            420                 425                 430
```

```
Glu Ala Thr Gly Ala Asn Ile Val Arg Val Ala Leu Glu Arg Pro Leu
        435                 440                 445

Lys Gln Ile Ala Phe Asn Gly Gly Leu Glu Pro Gly Val Val Ala Glu
    450                 455                 460

Lys Val Arg Asn Ser Pro Ala Gly Thr Gly Leu Asn Ala Ala Thr Gly
465                 470                 475                 480

Glu Tyr Glu Asp Leu Leu Lys Ala Gly Ile Ala Asp Pro Val Lys Val
                485                 490                 495

Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Ile Ala Gly Leu Phe Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Ala Asp Lys Pro Glu Lys Ala Ala Ala Pro
        515                 520                 525

Ala Gly Asp Pro Thr Gly Gly Met Gly Gly Met Asp Phe
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 540 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Ala Lys Thr Ile Ala Tyr Asp Glu Glu Ala Arg Arg Gly Leu Glu
1               5                   10                  15

Arg Gly Leu Asn Ala Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Lys Trp Gly Ala Pro Thr Ile
        35                  40                  45

Thr Asn Asp Gly Val Ser Ile Ala Lys Glu Ile Glu Leu Glu Asp Pro
    50                  55                  60

Tyr Glu Lys Ile Gly Ala Glu Leu Val Lys Glu Val Ala Lys Lys Thr
65                  70                  75                  80

Asp Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
                85                  90                  95

Ala Leu Val Arg Glu Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro
```

```
                100                 105                 110

Leu Gly Leu Lys Arg Gly Ile Glu Lys Ala Val Glu Lys Val Thr Glu
        115                 120                 125

Thr Leu Leu Lys Gly Ala Lys Glu Val Glu Thr Lys Glu Gln Ile Ala
    130                 135                 140

Ala Thr Ala Ala Ile Ser Ala Gly Asp Gln Ser Ile Gly Asp Leu Ile
145                 150                 155                 160

Ala Glu Ala Met Asp Lys Val Gly Asn Glu Gly Val Ile Thr Val Glu
                165                 170                 175

Glu Ser Asn Thr Phe Gly Leu Gln Leu Glu Leu Thr Glu Gly Met Arg
            180                 185                 190

Phe Asp Lys Gly Tyr Ile Ser Gly Tyr Phe Val Thr Asp Pro Glu Arg
        195                 200                 205

Gln Glu Ala Val Leu Glu Asp Pro Tyr Ile Leu Leu Val Ser Ser Lys
    210                 215                 220

Val Ser Thr Val Lys Asp Leu Leu Pro Leu Leu Glu Lys Val Ile Gly
225                 230                 235                 240

Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly Glu Ala
                245                 250                 255

Leu Ser Thr Leu Val Val Asn Lys Ile Arg Gly Thr Phe Lys Ser Val
            260                 265                 270

Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Gln
        275                 280                 285

Asp Met Ala Ile Leu Thr Gly Gly Gln Val Ile Ser Glu Glu Val Gly
    290                 295                 300

Leu Thr Leu Glu Asn Ala Asp Leu Ser Leu Leu Gly Lys Ala Arg Lys
305                 310                 315                 320

Val Val Val Thr Lys Asp Glu Thr Thr Ile Val Glu Gly Ala Gly Asp
                325                 330                 335

Thr Asp Ala Ile Ala Gly Arg Val Ala Gln Ile Arg Gln Glu Ile Glu
            340                 345                 350

Asn Ser Asp Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Ala Gly Ala Ala Thr Glu
    370                 375                 380

Val Glu Leu Lys Glu Arg Lys His Arg Ile Glu Asp Ala Val Arg Asn
385                 390                 395                 400
```

```
Ala Lys Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Val Thr
                405                 410                 415

Leu Leu Gln Ala Ala Pro Thr Leu Asp Glu Leu Lys Leu Glu Gly Asp
            420                 425                 430

Glu Ala Thr Gly Ala Asn Ile Val Lys Val Ala Leu Glu Ala Pro Leu
        435                 440                 445

Lys Gln Ile Ala Phe Asn Ser Gly Leu Glu Pro Gly Val Val Ala Glu
    450                 455                 460

Lys Val Arg Asn Leu Pro Ala Gly His Gly Leu Asn Ala Gln Thr Gly
465                 470                 475                 480

Val Tyr Glu Asp Leu Leu Ala Ala Gly Val Ala Asp Pro Val Lys Val
                485                 490                 495

Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Ile Ala Gly Leu Phe Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Ala Asp Lys Pro Glu Lys Glu Lys Ala Ser
        515                 520                 525

Val Pro Gly Gly Gly Asp Met Gly Gly Met Asp Phe
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 537 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Met Ser Lys Leu Ile Glu Tyr Asp Glu Thr Ala Arg His Ala Met Glu
1               5                   10                  15

Val Gly Met Asn Lys Leu Ala Asp Thr Val Arg Val Thr Leu Gly Pro
            20                  25                  30

Arg Gly Arg His Val Val Leu Ala Lys Ala Phe Gly Gly Pro Thr Ile
        35                  40                  45

Thr Asn Asp Gly Val Thr Val Ala Arg Glu Ile Asp Leu Glu Asp Pro
    50                  55                  60

Phe Glu Asn Leu Gly Ala Gln Leu Val Lys Ser Val Ala Thr Lys Thr
```

65 70 75 80

Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
85 90 95

Ala Leu Val Lys Gly Gly Leu Arg Met Val Ala Ala Gly Ala Asn Pro
100 105 110

Val Ala Leu Gly Ala Gly Ile Ser Lys Ala Ala Asp Ala Val Ser Glu
115 120 125

Ala Leu Leu Ala Val Ala Thr Pro Val Ala Gly Lys Asp Ala Ile Thr
130 135 140

Gln Val Ala Thr Val Ser Ser Arg Asp Glu Gln Ile Gly Ala Leu Val
145 150 155 160

Gly Glu Gly Met Asn Lys Val Gly Thr Asp Gly Val Val Ser Val Glu
165 170 175

Glu Ser Ser Thr Leu Asp Thr Glu Leu Glu Phe Thr Glu Gly Val Gly
180 185 190

Phe Asp Lys Gly Phe Leu Ser Ala Tyr Phe Val Thr Asp Phe Asp Ser
195 200 205

Gln Gln Ala Val Leu Asp Asp Pro Leu Val Leu Leu His Gln Glu Lys
210 215 220

Ile Ser Ser Leu Pro Glu Leu Leu Pro Met Leu Glu Lys Val Thr Glu
225 230 235 240

Ser Gly Lys Pro Leu Leu Ile Val Ala Glu Asp Leu Glu Gly Glu Ala
245 250 255

Leu Ala Thr Leu Val Val Asn Ser Ile Arg Lys Thr Leu Lys Ala Val
260 265 270

Ala Val Lys Ser Pro Phe Phe Gly Asp Arg Arg Lys Ala Phe Leu Glu
275 280 285

Asp Leu Ala Ile Val Thr Gly Gly Gln Val Val Asn Pro Glu Thr Gly
290 295 300

Leu Val Leu Arg Glu Val Gly Thr Asp Val Leu Gly Ser Ala Arg Arg
305 310 315 320

Val Val Val Ser Lys Asp Asp Thr Ile Ile Val Asp Gly Gly Gly Ser
325 330 335

Asn Asp Ala Val Ala Lys Arg Val Asn Gln Leu Arg Ala Glu Ile Glu
340 345 350

Val Ser Asp Ser Glu Trp Asp Arg Glu Lys Leu Gln Glu Arg Val Ala
355 360 365

```
Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Val Thr Glu
    370                 375                 380

Thr Ala Leu Lys Lys Arg Lys Glu Ser Val Glu Asp Ala Val Ala Ala
385                 390                 395                 400

Ala Lys Ala Ser Ile Glu Glu Gly Ile Ile Ala Gly Gly Gly Ser Ala
                405                 410                 415

Leu Val Gln Cys Gly Ala Ala Leu Lys Gln Leu Arg Thr Ser Leu Thr
            420                 425                 430

Gly Asp Glu Ala Leu Gly Ile Asp Val Phe Phe Glu Ala Leu Lys Ala
        435                 440                 445

Pro Leu Tyr Trp Ile Ala Thr Asn Ala Gly Leu Asp Gly Ala Val Val
    450                 455                 460

Val Asp Lys Val Ser Gly Leu Pro Ala Gly His Gly Leu Asn Ala Ser
465                 470                 475                 480

Thr Leu Gly Tyr Gly Asp Leu Val Ala Asp Gly Val Val Asp Pro Val
                485                 490                 495

Lys Val Thr Arg Ser Ala Val Leu Asn Ala Ala Ser Val Ala Arg Met
            500                 505                 510

Met Leu Thr Thr Glu Thr Ala Val Val Asp Lys Pro Ala Lys Thr Glu
        515                 520                 525

Glu His Asp His His Gly His Ala His
    530                 535
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 541 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Met Ala Lys Thr Ile Ala Tyr Asp Glu Glu Ala Arg Arg Gly Leu Glu
1               5                   10                  15

Arg Gly Leu Asn Ser Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Lys Trp Gly Ala Pro Thr Ile
```

```
        35                  40                  45

Thr Asn Asp Gly Val Ser Ile Ala Lys Glu Ile Glu Leu Glu Asp Pro
    50                  55                  60

Tyr Glu Lys Ile Gly Ala Glu Leu Val Lys Glu Val Ala Lys Lys Thr
65                  70                  75                  80

Asp Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
                85                  90                  95

Ala Leu Val Lys Glu Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro
            100                 105                 110

Leu Gly Leu Lys Arg Gly Ile Glu Lys Ala Val Asp Lys Val Thr Glu
        115                 120                 125

Thr Leu Leu Lys Asp Ala Lys Glu Val Glu Thr Lys Glu Gln Ile Ala
    130                 135                 140

Ala Thr Ala Ala Ile Ser Ala Gly Asp Gln Ser Ile Gly Asp Leu Ile
145                 150                 155                 160

Ala Glu Ala Met Asp Lys Val Gly Met Glu Gly Val Ile Thr Val Glu
                165                 170                 175

Glu Ser Asn Thr Phe Gly Leu Gln Leu Glu Leu Thr Glu Gly Met Arg
            180                 185                 190

Phe Asp Lys Gly Tyr Ile Ser Gly Tyr Phe Val Thr Asp Ala Glu Arg
        195                 200                 205

Gln Glu Ala Val Leu Glu Glu Pro Tyr Ile Leu Leu Val Ser Ser Lys
    210                 215                 220

Val Ser Thr Val Lys Asp Leu Leu Pro Leu Leu Glu Lys Val Ile Gln
225                 230                 235                 240

Ala Gly Lys Ser Leu Leu Ile Ile Ala Glu Asp Val Glu Gly Glu Ala
                245                 250                 255

Leu Ser Thr Leu Val Val Asn Lys Ile Arg Gly Thr Phe Lys Ser Val
            260                 265                 270

Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Gln
        275                 280                 285

Asp Met Ala Ile Leu Thr Gly Ala Gln Val Ile Ser Glu Glu Val Gly
    290                 295                 300

Leu Thr Leu Glu Asn Thr Asp Leu Ser Leu Leu Gly Lys Ala Arg Lys
305                 310                 315                 320

Val Val Met Thr Lys Asp Glu Thr Thr Ile Val Glu Gly Ala Gly Asp
                325                 330                 335
```

```
Thr Asp Ala Ile Ala Gly Arg Val Ala Gln Ile Arg Thr Glu Ile Glu
            340                 345                 350

Asn Ser Asp Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Ala Gly Ala Ala Thr Glu
    370                 375                 380

Val Glu Leu Lys Glu Arg Lys His Arg Ile Glu Asp Ala Val Arg Asn
385                 390                 395                 400

Ala Lys Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Val Thr
                405                 410                 415

Leu Leu Gln Ala Ala Pro Ala Leu Asp Lys Leu Lys Leu Thr Gly Asp
            420                 425                 430

Glu Ala Thr Gly Ala Asn Ile Val Lys Val Ala Leu Glu Ala Pro Leu
        435                 440                 445

Lys Gln Ile Ala Phe Asn Ser Gly Met Glu Pro Gly Val Val Ala Glu
    450                 455                 460

Lys Val Arg Asn Leu Ser Val Gly His Gly Leu Asn Ala Ala Thr Gly
465                 470                 475                 480

Glu Tyr Glu Asp Leu Leu Lys Ala Gly Val Ala Asp Pro Val Lys Val
                485                 490                 495

Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Ile Ala Gly Leu Phe Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Ala Asp Lys Pro Glu Lys Thr Ala Ala Pro
        515                 520                 525

Ala Ser Asp Pro Thr Gly Gly Met Gly Gly Met Asp Phe
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 539 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Met Ser Lys Leu Ile Glu Tyr Asp Glu Thr Ala Arg Arg Ala Met Glu
```

```
1               5                   10                  15

Val Gly Met Asp Lys Leu Ala Asp Thr Val Arg Val Thr Leu Gly Pro
            20                  25                  30

Arg Gly Arg His Val Val Leu Ala Lys Ala Phe Gly Gly Pro Thr Val
        35                  40                  45

Thr Asn Asp Gly Val Thr Val Ala Arg Glu Ile Glu Leu Glu Asp Pro
    50                  55                  60

Phe Glu Asp Leu Gly Ala Gln Leu Val Lys Ser Val Ala Thr Lys Thr
65                  70                  75                  80

Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Ile Leu Ala Gln
                85                  90                  95

Ala Leu Ile Lys Gly Gly Leu Arg Leu Val Ala Ala Gly Val Asn Pro
            100                 105                 110

Ile Ala Leu Gly Val Gly Ile Gly Lys Ala Ala Asp Ala Val Ser Glu
        115                 120                 125

Ala Leu Leu Ala Ser Ala Thr Pro Val Ser Gly Lys Thr Gly Ile Ala
    130                 135                 140

Gln Val Ala Thr Val Ser Ser Arg Asp Glu Gln Ile Gly Asp Leu Val
145                 150                 155                 160

Gly Glu Ala Met Ser Lys Val Gly His Asp Gly Val Val Ser Val Glu
                165                 170                 175

Glu Ser Ser Thr Leu Gly Thr Glu Leu Glu Phe Thr Glu Gly Ile Gly
            180                 185                 190

Phe Asp Lys Gly Phe Leu Ser Ala Tyr Phe Val Thr Asp Phe Asp Asn
        195                 200                 205

Gln Gln Ala Val Leu Glu Asp Ala Leu Ile Leu Leu His Gln Asp Lys
    210                 215                 220

Ile Ser Ser Leu Pro Asp Leu Leu Pro Leu Leu Glu Lys Val Ala Gly
225                 230                 235                 240

Thr Gly Lys Pro Leu Leu Ile Val Ala Glu Asp Val Glu Gly Glu Ala
                245                 250                 255

Leu Ala Thr Leu Val Val Asn Ala Ile Arg Lys Thr Leu Lys Ala Val
            260                 265                 270

Ala Val Lys Gly Pro Tyr Phe Gly Asp Arg Arg Lys Ala Phe Leu Glu
        275                 280                 285

Asp Leu Ala Val Val Thr Gly Gly Gln Val Val Asn Pro Asp Ala Gly
    290                 295                 300
```

```
Met Val Leu Arg Glu Val Gly Leu Glu Val Leu Gly Ser Ala Arg Arg
305                 310                 315                 320

Val Val Val Ser Lys Asp Asp Thr Val Ile Val Asp Gly Gly Gly Thr
                325                 330                 335

Ala Glu Ala Val Ala Asn Arg Ala Lys His Leu Arg Ala Glu Ile Asp
            340                 345                 350

Lys Ser Asp Ser Asp Trp Asp Arg Glu Lys Leu Gly Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu
    370                 375                 380

Thr Ala Leu Lys Glu Arg Lys Glu Ser Val Glu Asp Ala Val Ala Ala
385                 390                 395                 400

Ala Lys Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly Ala Ser
                405                 410                 415

Leu Ile His Gln Ala Arg Lys Ala Leu Thr Glu Leu Arg Ala Ser Leu
            420                 425                 430

Thr Gly Asp Glu Val Leu Gly Val Asp Val Phe Ser Glu Ala Leu Ala
        435                 440                 445

Ala Pro Leu Phe Trp Ile Ala Ala Asn Ala Gly Leu Asp Gly Ser Val
    450                 455                 460

Val Val Lys Lys Val Ser Glu Leu Pro Ala Gly His Gly Leu Asn Val
465                 470                 475                 480

Asn Thr Leu Ser Tyr Gly Asp Leu Ala Ala Asp Gly Val Ile Asp Pro
                485                 490                 495

Val Lys Val Thr Arg Ser Ala Val Leu Asn Ala Ser Ser Val Ala Arg
            500                 505                 510

Met Val Leu Thr Thr Glu Thr Val Val Val Asp Lys Pro Ala Lys Ala
        515                 520                 525

Glu Asp His Asp His His His Gly His Ala His
    530                 535
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 545 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Ala Ala Lys Asp Val Gln Phe Gly Asn Glu Val Arg Gln Lys Met
1               5                  10                  15

Val Asn Gly Val Asn Ile Leu Ala Asn Ala Val Arg Val Thr Leu Gly
            20                  25                  30

Pro Lys Gly Arg Asn Val Val Val Asp Arg Ala Phe Gly Gly Pro His
        35                  40                  45

Ile Thr Lys Asp Gly Val Thr Val Ala Lys Glu Ile Glu Leu Lys Asp
    50                  55                  60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Ser Lys
65                  70                  75                  80

Thr Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
                85                  90                  95

Gln Ser Ile Val Ala Glu Gly Met Lys Tyr Val Thr Ala Gly Met Asn
            100                 105                 110

Pro Thr Asp Leu Lys Arg Gly Ile Asp Lys Ala Val Ala Ala Leu Val
        115                 120                 125

Glu Glu Leu Lys Asn Ile Ala Lys Pro Cys Asp Thr Ser Lys Glu Ile
    130                 135                 140

Ala Gln Val Gly Ser Ile Ser Ala Asn Ser Asp Glu Gln Val Gly Ala
145                 150                 155                 160

Ile Ile Ala Glu Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr
                165                 170                 175

Val Glu Asp Gly Lys Ser Leu Glu Asn Glu Leu Asp Val Val Glu Gly
            180                 185                 190

Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Asn Asp Ala
        195                 200                 205

Glu Lys Gln Ile Ala Gly Leu Asp Asn Pro Phe Val Leu Leu Phe Asp
    210                 215                 220

Lys Lys Ile Ser Asn Ile Arg Asp Leu Leu Pro Val Leu Glu Gln Val
225                 230                 235                 240

Ala Lys Ala Ser Arg Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly
                245                 250                 255

Glu Ala Leu Ala Thr Leu Val Val Asn Asn Ile Arg Gly Ile Leu Lys
            260                 265                 270
```

```
Thr Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met
        275                 280                 285

Leu Gln Asp Ile Ala Ile Leu Thr Gly Gly Thr Val Ile Ser Glu Glu
    290                 295                 300

Val Gly Leu Ser Leu Glu Lys Ala Thr Leu Asp Asp Leu Gly Gln Ala
305                 310                 315                 320

Lys Arg Ile Glu Ile Gly Lys Glu Asn Thr Thr Ile Ile Asp Gly Phe
                325                 330                 335

Gly Asp Ala Ala Gln Ile Glu Ala Arg Val Ala Glu Ile Arg Gln Gln
            340                 345                 350

Ile Glu Thr Ala Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu Arg
        355                 360                 365

Val Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala
    370                 375                 380

Thr Glu Val Glu Met Lys Glu Lys Lys Asp Arg Val Glu Asp Ala Leu
385                 390                 395                 400

His Ala Thr Arg Ala Ala Val Glu Glu Gly Val Val Ala Gly Gly Gly
                405                 410                 415

Val Ala Leu Leu Arg Ala Arg Ala Ala Leu Glu Asn Leu His Thr Gly
            420                 425                 430

Asn Ala Asp Gln Asp Ala Gly Val Gln Ile Val Leu Arg Ala Val Glu
        435                 440                 445

Ser Pro Leu Arg Gln Ile Val Ala Asn Ala Gly Gly Glu Pro Ser Val
    450                 455                 460

Val Val Asn Lys Val Leu Glu Gly Lys Gly Asn Tyr Gly Tyr Asn Ala
465                 470                 475                 480

Gly Ser Gly Glu Tyr Gly Asp Met Ile Glu Met Gly Val Leu Asp Pro
                485                 490                 495

Ala Lys Val Thr Arg Ser Ala Leu Gln His Ala Ala Ser Ile Ala Gly
            500                 505                 510

Leu Met Leu Thr Thr Asp Cys Met Ile Ala Glu Ile Pro Glu Glu Lys
        515                 520                 525

Pro Ala Met Pro Asp Met Gly Gly Met Gly Gly Met Gly Gly Met Met
    530                 535                 540

Glx
545
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 539 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Met Val Lys Gln Leu Lys Phe Ser Glu Asp Ala Arg Gln Ala Met Leu
1               5                  10                  15

Arg Gly Val Asp Gln Leu Ala Asn Ala Val Lys Val Thr Ile Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Asp Lys Glu Phe Thr Ala Pro Leu Ile
        35                  40                  45

Thr Asn Asp Gly Val Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp Pro
    50                  55                  60

Tyr Glu Asn Met Gly Ala Lys Leu Val Gln Glu Val Ala Asn Lys Thr
65                  70                  75                  80

Asn Glu Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
                85                  90                  95

Ala Met Ile Gln Glu Gly Leu Lys Asn Val Thr Ser Gly Ala Asn Pro
            100                 105                 110

Val Gly Leu Arg Gln Gly Ile Asp Lys Ala Val Lys Val Ala Val Glu
        115                 120                 125

Ala Leu His Glu Asn Ser Gln Lys Val Glu Asn Lys Asn Glu Ile Ala
    130                 135                 140

Gln Val Gly Ala Ile Ser Ala Ala Asp Glu Glu Ile Gly Arg Tyr Ile
145                 150                 155                 160

Ser Glu Ala Thr Glu Lys Val Gly Asn Asp Gly Val Ile Thr Ile Ile
                165                 170                 175

Thr Ile Glu Glu Ser Asn Arg Leu Asn Thr Glu Leu Glu Leu Gly Met
            180                 185                 190

Gln Phe Asp Arg Gly Tyr Gln Ser Pro Tyr Met Val Thr Asp Ser Asp
        195                 200                 205

Lys Met Val Ala Glu Leu Glu Arg Pro Tyr Ile Leu Val Thr Asp Lys
    210                 215                 220
```

```
Lys Ile Ser Ser Phe Gln Asp Ile Leu Pro Leu Leu Glu Gln Val Val
225                 230                 235                 240

Gln Ser Asn Arg Pro Ile Leu Ile Val Ala Asp Glu Val Glu Gly Asp
                245                 250                 255

Ala Leu Thr Asn Ile Val Leu Asn Arg Met Arg Gly Thr Phe Thr Ala
            260                 265                 270

Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu
        275                 280                 285

Glu Asp Leu Ala Ile Leu Thr Gly Ala Gln Val Ile Thr Asp Asp Leu
    290                 295                 300

Gly Leu Asp Leu Lys Asp Ala Ser Ile Asp Met Leu Gly Thr Ala Ser
305                 310                 315                 320

Lys Val Glu Val Thr Lys Asp Asn Thr Thr Val Val Asp Gly Asp Gly
                325                 330                 335

Asp Glu Asn Ser Ile Asp Ala Arg Val Ser Gln Leu Lys Ser Gln Ile
            340                 345                 350

Glu Glu Thr Glu Ser Asp Phe Asp Arg Glu Lys Leu Gln Glu Arg Leu
        355                 360                 365

Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Ser
    370                 375                 380

Glu Thr Glu Leu Lys Glu Arg Lys Leu Arg Ile Glu Asp Ala Leu Asn
385                 390                 395                 400

Ser Thr Arg Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Thr
                405                 410                 415

Ala Leu Val Asn Val Tyr Gln Lys Val Ser Glu Asn Glu Ala Glu Gly
            420                 425                 430

Asp Ile Glu Thr Gly Val Asn Ile Val Leu Lys Ala Leu Thr Ala Pro
        435                 440                 445

Val Arg Gln Ile Ala Glu Asn Ala Gly Leu Glu Gly Ser Val Ile Val
    450                 455                 460

Glu Arg Leu Lys Asn Ala Glu Pro Gly Val Gly Phe Asn Gly Ala Thr
465                 470                 475                 480

Asn Glu Trp Val Asn Met Leu Arg Arg Gly Ile Val Asp Pro Thr Lys
                485                 490                 495

Val Thr Arg Ser Ala Leu Gln His Ala Ala Ser Val Ala Ala Met Phe
            500                 505                 510
```

```
Leu Thr Thr Glu Ala Val Val Ala Ser Ile Pro Glu Lys Asn Asn Asp
        515                 520                 525

Gln Pro Asn Met Gly Gly Met Pro Gly Met Met
    530                 535
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 541 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Met Ala Lys Ser Ile Ile Tyr Asn Asp Glu Ala Arg Arg Ala Leu Glu
1               5                  10                  15

Arg Gly Met Asp Ile Leu Ala Glu Ala Val Ala Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Lys Phe Gly Ser Pro Gln Ile
        35                  40                  45

Ile Asn Asp Gly Ile Thr Ile Ala Lys Glu Ile Glu Leu Glu Asp His
    50                  55                  60

Val Glu Asn Thr Gly Val Ser Leu Ile Arg Gln Ala Ala Ser Lys Thr
65                  70                  75                  80

Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala His
                85                  90                  95

Ala Ile Val Lys Glu Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro
            100                 105                 110

Ile Ser Leu Lys Arg Gly Ile Asp Lys Ala Thr Asp Phe Leu Val Ala
        115                 120                 125

Arg Ile Lys Glu His Ala Gln Pro Val Gly Asp Ser Lys Ala Ile Ala
    130                 135                 140

Gln Val Gly Ala Ile Ser Ala Gly Asn Asp Glu Glu Val Gly Gln Met
145                 150                 155                 160

Ile Ala Asn Ala Met Asp Lys Val Gly Gln Glu Gly Val Ile Ser Leu
                165                 170                 175

Glu Glu Gly Lys Ser Met Thr Thr Glu Leu Glu Ile Thr Glu Gly Met
            180                 185                 190
```

```
Arg Phe Asp Lys Gly Tyr Ile Ser Pro Tyr Phe Val Thr Asp Ala Glu
        195                 200                 205

Arg Met Glu Ala Val Leu Glu Asp Pro Arg Ile Leu Ile Thr Asp Lys
    210                 215                 220

Lys Ile Asn Leu Val Gln Asp Leu Val Pro Ile Leu Glu Gln Val Ala
225                 230                 235                 240

Arg Gln Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Ile Glu Lys Glu
                245                 250                 255

Ala Leu Ala Thr Leu Val Val Asn Arg Leu Arg Gly Val Leu Asn Val
            260                 265                 270

Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Gln Met Leu
        275                 280                 285

Glu Asp Ile Ala Thr Leu Thr Gly Gly Gln Val Ile Ser Glu Asp Ala
    290                 295                 300

Gly Leu Lys Leu Glu Ser Ala Thr Val Asp Ser Leu Gly Ser Ala Arg
305                 310                 315                 320

Arg Ile Asn Ile Thr Lys Asp Asn Thr Thr Ile Val Ala Glu Gly Asn
                325                 330                 335

Glu Ala Ala Val Lys Ser Arg Cys Glu Gln Ile Arg Arg Gln Ile Glu
            340                 345                 350

Glu Thr Asp Ser Ser Tyr Asp Lys Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu
    370                 375                 380

Thr Glu Met Lys Asp Arg Lys Leu Arg Leu Glu Asp Ala Ile Asn Ala
385                 390                 395                 400

Thr Lys Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly Thr Thr
                405                 410                 415

Leu Ala His Leu Ala Pro Gln Leu Glu Asp Trp Ala Thr Gly Asn Leu
            420                 425                 430

Lys Asp Glu Glu Leu Thr Gly Ala Leu Ile Val Ala Arg Ala Leu Pro
        435                 440                 445

Ala Pro Leu Lys Arg Ile Ala Glu Asn Ala Gly Gln Asn Gly Ala Val
    450                 455                 460

Ile Ser Glu Arg Val Lys Glu Lys Glu Phe Asn Val Gly Tyr Asn Ala
465                 470                 475                 480
```

```
Ala Ser Leu Glu Tyr Val Asp Met Leu Ala Ala Gly Ile Val Asp Pro
                485                 490                 495

Ala Lys Val Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Ile Ala Gly
            500                 505                 510

Met Val Leu Thr Thr Glu Cys Ile Val Val Asp Lys Pro Glu Lys Glu
        515                 520                 525

Lys Ala Pro Ala Gly Ala Pro Gly Gly Asp Phe Asp Tyr
    530                 535                 540
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 552 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Met Ser Lys Leu Ile Ser Phe Lys Asp Glu Ser Arg Arg Ser Leu Glu
1               5                   10                  15

Ala Gly Ile Asn Ala Leu Ala Asp Ala Val Arg Ile Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Leu Leu Glu Lys Gln Tyr Gly Ala Pro Gln Ile
        35                  40                  45

Val Asn Asp Gly Ile Thr Val Ala Lys Glu Ile Glu Leu Ser Asn Pro
    50                  55                  60

Glu Glu Asn Ala Gly Ala Lys Leu Ile Gln Glu Val Ala Ser Lys Thr
65                  70                  75                  80

Lys Glu Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Ile Ile Ala Gln
                85                  90                  95

Ala Leu Val Arg Glu Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro
            100                 105                 110

Val Ala Leu Arg Arg Gly Ile Glu Lys Val Thr Thr Phe Leu Val Gln
        115                 120                 125

Glu Ile Glu Ala Val Ala Lys Pro Val Glu Gly Ser Ala Ile Ala Gln
    130                 135                 140

Val Ala Thr Val Ser Ser Gly Asn Asp Pro Glu Val Gly Ala Met Ile
145                 150                 155                 160
```

```
Ala Asp Ala Met Asp Lys Val Thr Lys Asp Gly Val Ile Thr Val Glu
                165                 170                 175

Glu Ser Lys Ser Leu Asn Thr Glu Leu Glu Val Val Glu Gly Met Gln
            180                 185                 190

Ile Asp Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Thr Asp Ser Asp Arg
        195                 200                 205

Gln Leu Val Glu Phe Asp Asn Pro Leu Ile Leu Ile Thr Asp Lys Lys
    210                 215                 220

Ile Ser Ala Ile Ala Glu Leu Val Pro Val Leu Glu Ala Val Ala Arg
225                 230                 235                 240

Ala Gly Arg Pro Leu Leu Ile Ile Ala Glu Asp Ile Glu Gly Glu Ala
                245                 250                 255

Leu Ala Thr Leu Val Val Asn Lys Ala Arg Gly Val Leu Asn Val Ala
            260                 265                 270

Ala Ile Lys Ala Pro Ala Phe Gly Asp Arg Arg Lys Ala Val Leu Gln
        275                 280                 285

Asp Ile Ala Ile Leu Thr Gly Gly Ser Val Ile Ser Glu Asp Ile Gly
    290                 295                 300

Leu Ser Leu Asp Thr Val Ser Leu Asp Gln Leu Gly Gln Ala Val Lys
305                 310                 315                 320

Ala Thr Leu Glu Lys Asp Asn Thr Ile Leu Val Ala Gly Ala Asp Lys
                325                 330                 335

Arg Ala Ser Ala Gly Val Lys Glu Arg Ile Glu Gln Leu Arg Lys Glu
            340                 345                 350

Tyr Ala Ala Ser Asp Ser Asp Tyr Asp Lys Glu Lys Ile Gln Glu Arg
        355                 360                 365

Ile Ala Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala
    370                 375                 380

Thr Glu Thr Glu Leu Lys Asp Arg Lys Leu Arg Ile Glu Asp Ala Leu
385                 390                 395                 400

Asn Ala Thr Lys Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly
                405                 410                 415

Thr Thr Leu Ile Arg Leu Ala Gly Lys Ile Glu Ser Phe Lys Ala Gln
            420                 425                 430

Leu Ser Asn Asp Glu Glu Arg Val Ala Ala Asp Ile Ile Ala Lys Ala
        435                 440                 445
```

```
Leu Glu Ala Pro Leu His Gln Leu Ala Ser Asn Ala Gly Val Glu Gly
    450                 455                 460

Ser Val Ile Val Glu Lys Val Lys Glu Ala Thr Gly Asn Gln Gly Tyr
465                 470                 475                 480

Asn Val Ile Thr Gly Lys Ile Glu Asp Leu Ile Ala Ala Gly Ile Ile
                485                 490                 495

Asp Pro Ala Lys Val Val Arg Ser Ala Leu Gln Asn Ala Ala Ser Ile
            500                 505                 510

Ala Gly Met Val Leu Thr Thr Glu Ala Leu Val Val Glu Lys Pro Glu
        515                 520                 525

Pro Ala Ala Pro Ala Met Pro Asp Met Gly Gly Met Gly Gly Met Gly
    530                 535                 540

Gly Met Gly Gly Met Gly Met Met
545                 550
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 539 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Met Ala Lys Thr Ile Ala Phe Asp Lys Lys Ala Arg Arg Gly Leu Glu
1               5                   10                  15

Arg Gly Leu Asn Ala Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Lys Trp Gly Ala Pro Thr Ile
        35                  40                  45

Thr Asn Asp Gly Val Ser Ile Ala Lys Glu Ile Glu Leu Glu Asp Pro
    50                  55                  60

Tyr Glu Lys Ile Gly Ala Glu Leu Val Lys Glu Val Ala Lys Lys Thr
65                  70                  75                  80

Asp Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
                85                  90                  95

Ala Leu Val Arg Glu Gly Leu Arg Asn Val Ala Ala Gly Ala Asn Pro
            100                 105                 110
```

```
Leu Gly Leu Lys Arg Gly Ile Glu Lys Ala Val Glu Ala Val Thr Glu
        115                 120                 125

His Leu Leu Lys Ala Ala Lys Glu Val Glu Thr Lys Asp Gln Ile Ala
    130                 135                 140

Ala Thr Ala Gly Ile Ser Ala Gly Asp Pro Ala Ile Gly Glu Leu Ile
145                 150                 155                 160

Ala Glu Ala Met Asp Lys Val Gly Lys Glu Gly Val Ile Thr Val Glu
                165                 170                 175

Glu Ser Asn Thr Phe Gly Leu Gln Leu Glu Leu Thr Glu Gly Met Arg
            180                 185                 190

Phe Asp Lys Gly Phe Ile Ser Gly Tyr Phe Ala Thr Asp Ala Glu Arg
        195                 200                 205

Gln Glu Ala Val Leu Glu Asp Pro Tyr Val Leu Leu Val Ser Gly Lys
    210                 215                 220

Ile Ser Thr Val Lys Asp Leu Leu Pro Leu Leu Glu Lys Val Ile Gln
225                 230                 235                 240

Ser Gly Lys Pro Leu Ala Ile Ile Ala Glu Asp Val Glu Gly Glu Ala
                245                 250                 255

Leu Val Thr Leu Ile Val Asn Lys Ile Arg Gly Thr Phe Lys Ser Val
            260                 265                 270

Ala Ile Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu Gln
        275                 280                 285

Asp Met Ala Ile Leu Thr Gly Gly Gln Val Ile Ser Glu Glu Ile Gly
    290                 295                 300

Leu Ser Leu Asp Thr Ala Gly Leu Glu Val Leu Gly Gln Ala Arg Gln
305                 310                 315                 320

Val Val Val Thr Lys Asp Glu Thr Thr Ile Val Asp Gly Ala Gly Ser
                325                 330                 335

Lys Glu Gln Ile Ala Gly Arg Val Ser Gln Ile Arg Ala Glu Ile Glu
            340                 345                 350

Asn Ser Asp Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Ala Gly Ala Ala Thr Glu
    370                 375                 380

Asp Leu Lys Glu Arg Lys His Arg Ile Glu Asp Ala Val Arg Asn Ala
385                 390                 395                 400
```

```
Lys Ala Ala Val Glu Glu Gly Ile Val Ala Gly Gly Gly Ser Ser Leu
                405                 410                 415

Ala Gln Ser Gly Thr Val Phe Asp Ser Xaa Ala Leu Glu Gly Asp Glu
            420                 425                 430

Ala Thr Gly Ala Asn Ile Val Lys Val Ala Leu Asp Ala Pro Val Lys
        435                 440                 445

Gln Ile Ala Val Asn Ala Gly Leu Glu Pro Gly Val Val Ala Glu Lys
    450                 455                 460

Val Arg Asn Ser Pro Ala Gly Thr Gly Leu Asn Ala Ala Thr Gly Val
465                 470                 475                 480

Tyr Glu Asp Leu Leu Ala Ala Gly Ile Asn Asp Pro Val Lys Val Thr
                485                 490                 495

Arg Ser Ala Leu Gln Asn Ala Ala Ser Ile Ala Ala Leu Phe Leu Thr
            500                 505                 510

Thr Glu Ala Val Val Ala Asp Lys Pro Glu Lys Ala Gly Ala Pro Val
        515                 520                 525

Asp Pro Thr Gly Gly Met Gly Gly Met Asp Phe
    530                 535
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 582 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
Met Val Ser Phe Leu Ser Ser Ser Val Ser Arg Leu Pro Leu Arg Ile
1               5                   10                  15

Ala Gly Arg Arg Ile Pro Gly Arg Phe Ala Val Pro Gln Val Arg Thr
            20                  25                  30

Tyr Ala Lys Asp Leu Lys Phe Gly Val Asp Ala Arg Ala Ser Leu Leu
        35                  40                  45

Thr Gly Val Asp Thr Leu Ala Arg Ala Val Ser Val Thr Leu Gly Pro
    50                  55                  60

Lys Gly Arg Asn Val Leu Ile Asp Gln Pro Phe Gly Ser Pro Lys Ile
65                  70                  75                  80
```

```
Thr Lys Asp Gly Val Thr Val Ala Arg Ser Val Ser Leu Lys Asp Lys
                85                  90                  95

Phe Glu Asn Leu Gly Ala Arg Leu Val Gln Asp Val Ala Ser Lys Thr
            100                 105                 110

Asn Glu Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Thr Arg
        115                 120                 125

Ala Ile Phe Ser Glu Thr Val Arg Asn Val Ala Ala Gly Cys Asn Pro
    130                 135                 140

Met Asp Leu Arg Arg Gly Ile Gln Leu Ala Val Asp Asn Val Val Glu
145                 150                 155                 160

Phe Leu Gln Ala Asn Lys Arg Asp Ile Thr Thr Ser Glu Glu Ile Ser
                165                 170                 175

Gln Val Ala Thr Ile Ser Ala Asn Gly Asp Thr His Ile Gly Glu Leu
            180                 185                 190

Leu Ala Lys Ala Met Glu Arg Val Gly Lys Glu Gly Val Ile Thr Val
        195                 200                 205

Lys Glu Gly Arg Thr Ile Ser Asp Glu Leu Glu Val Thr Glu Gly Met
    210                 215                 220

Lys Phe Asp Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Thr Asp Val Lys
225                 230                 235                 240

Ser Gln Lys Val Glu Phe Glu Asn Pro Leu Ile Leu Leu Ser Glu Lys
                245                 250                 255

Lys Val Ser Ala Val Gln Asp Ile Leu Pro Ser Leu Glu Leu Ala Ala
            260                 265                 270

Gln Gln Arg Arg Pro Leu Val Ile Ile Ala Glu Asp Val Asp Gly Glu
        275                 280                 285

Ala Leu Ala Ala Cys Ile Leu Asn Lys Leu Arg Gly Gln Leu Gln Val
    290                 295                 300

Val Ala Ile Lys Ala Pro Gly Phe Gly Asp Asn Arg Arg Asn Met Leu
305                 310                 315                 320

Gly Asp Leu Ala Val Leu Thr Asp Ser Ala Val Phe Asn Asp Glu Ile
                325                 330                 335

Asp Val Ser Ile Glu Lys Ala Gln Pro His His Leu Gly Ser Cys Gly
            340                 345                 350

Ser Val Thr Val Thr Lys Glu Asp Thr Ile Ile Met Lys Gly Ala Gly
        355                 360                 365
```

```
Asp His Val Lys Val Asn Asp Arg Cys Glu Gln Ile Arg Gly Val Met
    370                 375                 380

Ala Asp Pro Asn Leu Thr Glu Ser Glu Lys Glu Lys Leu Gln Glu Arg
385                 390                 395                 400

Leu Ala Lys Leu Ser Gly Gly Ile Ala Val Ile Lys Val Gly Ala Ser
                405                 410                 415

Ser Glu Val Glu Val Asn Glu Lys Lys Asp Arg Ile Val Asp Ala Leu
            420                 425                 430

Asn Ala Val Lys Ala Ala Val Ser Glu Gly Val Leu Pro Gly Ala Gly
        435                 440                 445

Thr Ser Phe Val Lys Ala Ser Leu Arg Leu Gly Asp Ile Pro Thr Asn
    450                 455                 460

Asn Phe Asp Gln Lys Leu Gly Val Glu Ile Val Arg Lys Ala Ile Thr
465                 470                 475                 480

Arg Pro Ala Gln Thr Ile Leu Glu Asn Ala Gly Leu Glu Gly Asn Leu
                485                 490                 495

Ile Val Gly Lys Leu Lys Glu Leu Tyr Gly Lys Glu Phe Asn Ile Gly
            500                 505                 510

Tyr Asp Ile Ala Lys Asp Arg Phe Val Asp Leu Asn Glu Ile Gly Val
        515                 520                 525

Leu Asp Pro Leu Lys Val Val Arg Thr Gly Leu Val Asp Ala Ser Gly
    530                 535                 540

Val Ala Ser Leu Met Gly Thr Thr Glu Cys Ala Ile Val Asp Ala Pro
545                 550                 555                 560

Glu Glu Ser Lys Ala Pro Ala Gly Pro Pro Gly Met Gly Gly Met Gly
                565                 570                 575

Gly Met Pro Gly Met Met
            580
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 572 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
Met Leu Arg Ser Ser Val Val Arg Ser Arg Ala Thr Leu Arg Pro Leu
1               5                  10                  15

Leu Arg Arg Ala Tyr Ser Ser His Lys Glu Leu Lys Phe Gly Val Glu
            20                  25                  30

Gly Arg Ala Ser Leu Leu Lys Gly Val Glu Thr Leu Ala Glu Ala Val
        35                  40                  45

Ala Ala Thr Leu Gly Pro Lys Gly Arg Asn Val Leu Ile Glu Gln Pro
    50                  55                  60

Phe Gly Pro Pro Lys Ile Thr Lys Asp Gly Val Thr Val Ala Lys Ser
65                  70                  75                  80

Ile Val Leu Lys Asp Lys Phe Glu Asn Met Gly Ala Lys Leu Leu Gln
                85                  90                  95

Glu Val Ala Ser Lys Thr Asn Glu Ala Ala Gly Asp Gly Thr Thr Ser
            100                 105                 110

Ala Thr Val Leu Gly Arg Ala Ile Phe Thr Glu Ser Val Lys Asn Val
        115                 120                 125

Ala Ala Gly Cys Asn Pro Met Asp Leu Arg Arg Gly Ser Gln Val Ala
    130                 135                 140

Val Glu Lys Val Ile Glu Phe Leu Ser Ala Asn Lys Lys Glu Ile Thr
145                 150                 155                 160

Thr Ser Glu Glu Ile Ala Gln Val Ala Thr Ile Ser Ala Asn Gly Asp
                165                 170                 175

Ser His Val Gly Lys Leu Leu Ala Ser Ala Met Glu Lys Val Gly Lys
            180                 185                 190

Glu Gly Val Ile Thr Ile Arg Glu Gly Arg Thr Leu Glu Asp Glu Leu
        195                 200                 205

Glu Val Thr Glu Gly Met Arg Phe Asp Arg Gly Phe Ile Ser Pro Tyr
    210                 215                 220

Phe Ile Thr Asp Pro Lys Ser Ser Lys Val Glu Phe Glu Lys Pro Leu
225                 230                 235                 240

Leu Leu Leu Ser Glu Lys Lys Ile Ser Ser Ile Gln Asp Ile Leu Pro
                245                 250                 255

Ala Leu Glu Ile Ser Asn Gln Ser Arg Arg Pro Leu Leu Ile Ile Ala
            260                 265                 270

Glu Asp Val Asp Gly Glu Ala Leu Ala Ala Cys Ile Leu Asn Lys Leu
        275                 280                 285
```

```
Arg Gly Gln Val Lys Val Cys Ala Val Lys Ala Pro Gly Phe Gly Asp
    290                 295                 300

Asn Arg Lys Asn Thr Ile Gly Asp Ile Ala Val Leu Thr Gly Gly Thr
305                 310                 315                 320

Val Phe Thr Glu Glu Leu Asp Leu Lys Pro Glu Gln Cys Thr Ile Glu
                325                 330                 335

Asn Leu Gly Ser Cys Asp Ser Ile Thr Val Thr Lys Glu Asp Thr Val
            340                 345                 350

Ile Leu Asn Gly Ser Gly Pro Lys Glu Ala Ile Gln Glu Arg Ile Glu
        355                 360                 365

Gln Ile Lys Gly Ser Ile Asp Ile Thr Thr Thr Asn Ser Tyr Glu Lys
    370                 375                 380

Glu Lys Leu Gln Glu Arg Leu Ala Lys Leu Ser Gly Gly Val Ala Val
385                 390                 395                 400

Ile Arg Val Gly Gly Ala Ser Glu Val Glu Val Gly Glu Lys Lys Asp
                405                 410                 415

Arg Tyr Asp Asp Ala Leu Asn Ala Thr Arg Ala Ala Val Glu Glu Gly
            420                 425                 430

Ile Leu Pro Gly Gly Gly Thr Ala Leu Val Lys Ala Ser Arg Val Leu
        435                 440                 445

Asp Glu Val Val Val Asp Asn Phe Asp Gln Lys Leu Gly Val Asp Ile
    450                 455                 460

Ile Arg Lys Ala Ile Thr Arg Pro Ala Lys Gln Ile Ile Glu Asn Ala
465                 470                 475                 480

Gly Glu Glu Gly Ser Val Ile Ile Gly Lys Leu Ile Asp Glu Tyr Gly
                485                 490                 495

Asp Asp Phe Ala Lys Gly Tyr Asp Ala Ser Lys Ser Glu Tyr Thr Asp
            500                 505                 510

Met Leu Ala Thr Gly Ile Ile Asp Pro Phe Lys Val Val Arg Ser Gly
        515                 520                 525

Leu Val Asp Ala Ser Gly Val Ala Ser Leu Leu Ala Thr Thr Glu Val
    530                 535                 540

Ala Ile Val Asp Ala Pro Glu Pro Pro Ala Ala Ala Gly Ala Gly Gly
545                 550                 555                 560

Met Pro Gly Gly Met Pro Gly Met Pro Gly Met Met
                565                 570
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 577 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Met Ile Ser Thr Leu Arg Gly Lys Ile Phe Asn Asn Gly Ser Asn Arg
1               5                  10                  15

Asn Lys Cys Val Ser Ile Leu Ser Asn Ile Gln Lys Arg Asn Ile Ser
            20                  25                  30

Lys Asp Ile Arg Phe Gly Ser Asp Ala Arg Thr Ala Met Leu Thr Gly
        35                  40                  45

Cys Asn Lys Leu Ala Asp Ala Val Ser Val Thr Leu Gly Pro Lys Gly
    50                  55                  60

Arg Asn Val Ile Ile Glu Gln Ser Phe Gly Ser Pro Lys Ile Thr Lys
65                  70                  75                  80

Asp Gly Val Thr Val Ala Lys Ser Ile Glu Phe Asn Asn Lys Leu Ala
                85                  90                  95

Asn Leu Gly Ala Gln Met Val Lys Gln Val Ala Ala Asn Thr Asn Gly
            100                 105                 110

Lys Ala Gly Asp Gly Thr Thr Thr Ala Thr Ile Leu Ala Arg Ser Ile
        115                 120                 125

Phe Gln Gln Gly Cys Lys Ala Val Asp Ser Gly Met Asn Pro Met Asp
    130                 135                 140

Leu Leu Arg Gly Ile Asn Lys Gly Val Glu Lys Val Leu Glu Tyr Leu
145                 150                 155                 160

Asn Ser Ile Lys Lys Asp Val Thr Thr Thr Glu Glu Ile Phe Asn Val
                165                 170                 175

Ala Ser Ile Ser Asn Gly Asp Lys Asn Ile Gly Gln Leu Ile Ala Asp
            180                 185                 190

Thr Met Lys Lys Val Gly Lys Glu Gly Thr Ile Thr Val Thr Glu Gly
        195                 200                 205

Lys Thr Leu Gln His Glu Leu Glu Ile Val Glu Gly Ile Lys Phe Asp
    210                 215                 220
```

```
Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Asn Asn Ser Gln Lys Val Glu
225                 230                 235                 240

Leu Asp Lys Pro Tyr Ile Leu Ile His Glu Lys Lys Ile Ser Thr Val
                245                 250                 255

Lys Ser Leu Leu Pro Val Leu Glu His Val Leu Gln Asn Gln Ser Ser
            260                 265                 270

Leu Leu Val Ile Ala Glu Asp Val Asp Ser Asp Ala Leu Ala Thr Leu
        275                 280                 285

Ile Val Asn Lys Leu Arg Leu Gly Leu Lys Ile Cys Ala Val Lys Ala
    290                 295                 300

Pro Gly Phe Gly Glu His Arg Lys Ala Leu Ile His Asp Ile Ala Val
305                 310                 315                 320

Met Thr Gly Ala Lys Val Ile Thr Glu Glu Thr Gly Leu Lys Leu Asp
                325                 330                 335

Asp Pro Gln Val Val Ser Tyr Leu Gly Lys Ala Lys Ser Ile Asn Val
            340                 345                 350

Thr Lys Asp Ser Thr Leu Ile Met Glu Gly Glu Gly Lys Lys Glu Glu
        355                 360                 365

Ile Asn Glu Arg Cys Glu Ser Ile Arg Asn Ala Ile Lys Met Asn Thr
    370                 375                 380

Ser Asp Tyr Glu Lys Glu Lys Leu Gln Glu Arg Leu Ala Lys Ile Thr
385                 390                 395                 400

Gly Gly Val Ala Leu Ile Lys Val Gly Gly Ile Ser Glu Val Glu Val
                405                 410                 415

Asn Glu Ile Lys Asp Arg Ile Gln Asp Ala Leu Cys Ala Thr Lys Ala
            420                 425                 430

Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly Ser Ala Leu Leu Phe
        435                 440                 445

Ala Ser Lys Glu Leu Asp Ser Val Gln Thr Asp Asn Tyr Asp Gln Arg
    450                 455                 460

Val Gly Val Asn Ile Ile Lys Asp Ala Cys Lys Ala Pro Ile Lys Gln
465                 470                 475                 480

Ile Ala Glu Asn Ala Gly His Glu Gly Ser Val Val Ala Gly Asn Ile
                485                 490                 495

Leu Lys Asp Lys Asn Ser Asn Ile Gly Phe Asn Ala Gln Glu Gly Lys
            500                 505                 510

Tyr Val Asp Met Ile Glu Ser Gly Ile Ile Asp Pro Thr Lys Val Val
```

```
        515                 520                 525

Lys Thr Ala Ile Ser Asp Ala Ala Ser Ile Ala Ser Leu Met Thr Thr
    530                 535                 540

Thr Glu Val Ala Ile Val Asp Phe Lys Asp Ser Lys Asn Glu Glu Ser
545                 550                 555                 560

Ser Gln His Met Asn Ser Val Asn Ser Met Gly Asp Met Gly Gly Met
                565                 570                 575

Tyr
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 550 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Met Thr Asn Val Val Val Ser Gly Glu Gln Leu Gln Gln Ala Phe Arg
1               5                   10                  15

Glu Val Ala Ala Val Ile Asp Ser Thr Val Ala Val Thr Ala Gly Pro
            20                  25                  30

Arg Gly Lys Thr Val Gly Ile Asn Lys Pro Tyr Gly Ala Pro Glu Ile
        35                  40                  45

Thr Lys Asp Gly Tyr Lys Val Met Lys Gly Ile Lys Pro Glu Lys Pro
    50                  55                  60

Leu Asn Ala Ala Ile Thr Ser Ile Phe Ala Gln Ser Cys Ser Gln Cys
65                  70                  75                  80

Asn Asp Lys Val Gly Asp Gly Thr Thr Thr Cys Ser Ile Leu Thr Ser
                85                  90                  95

Gly Met Ile Val Glu Ala Ser Lys Ser Ile Ala Ala Gly Asn Asp Arg
            100                 105                 110

Ile Ser Ile Lys Asn Gly Met Gln Lys Ala Lys Asp Val Val Leu Lys
        115                 120                 125

Glu Val Ala Ser Met Ala Arg Thr Ile Ser Leu Glu Lys Ile Asp Glu
    130                 135                 140
```

```
Val Ala Gln Val Ala Ile Ile Ser Ala Asn Gly Asp Arg Ser Ile Gly
145                 150                 155                 160

Ser Asn Ile Ala Asp Ala Val Lys Lys Val Gly Lys Glu Gly Val Ile
                165                 170                 175

Thr Val Glu Glu Ser Lys Gly Ser Lys Glu Leu Glu Val Glu Leu Thr
            180                 185                 190

Thr Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Ile Thr
        195                 200                 205

Asn Asn Glu Lys Met Ile Val Glu Leu Asp Asp Pro Tyr Leu Leu Ile
    210                 215                 220

Thr Glu Lys Lys Leu Asn Ile Ile Gln Pro Leu Leu Ser Ile Leu Glu
225                 230                 235                 240

Ala Val Val Lys Ser Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Ile
                245                 250                 255

Glu Gly Glu Ala Leu Ser Thr Leu Val Ile Asn Lys Leu Arg Gly Gly
            260                 265                 270

Leu Lys Val Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys
        275                 280                 285

Glu Met Leu Glu Asp Ile Ala Ala Leu Thr Asn Ala Lys Tyr Val Ile
    290                 295                 300

Lys Asp Glu Leu Gly Ile Lys Met Glu Asp Leu Thr Leu Glu Asp Leu
305                 310                 315                 320

Gly Ile Ala Lys Asn Val Lys Ile Thr Lys Asp Asn Thr Thr Ile Val
                325                 330                 335

Ser Glu Asn Arg Val Thr Asp Arg Val Lys Ala Arg Ile Glu Gln Ile
            340                 345                 350

Lys Ser Gln Ile Glu Ser Ser Thr Ser Asp Tyr Asp Lys Glu Lys Leu
        355                 360                 365

Arg Glu Arg Leu Ala Lys Leu Ser Gly Gly Val Ala Val Leu Lys Val
    370                 375                 380

Gly Gly Ala Thr Glu Leu Glu Val Lys Glu Arg Arg Asp Arg Val Glu
385                 390                 395                 400

Asp Gln Leu His Ala Thr Arg Ala Ala Ile Glu Glu Gly Ile Val Pro
                405                 410                 415

Gly Gly Gly Val Ala Leu Leu Tyr Ala Ser Ser Ala Leu Asp Lys Leu
            420                 425                 430

Lys Gly Ala Asp Asp Glu Glu Gln Ile Gly Ile Asn Ile Ile Lys Lys
```

435 440 445

Val Leu Ser Val Pro Ile Lys Arg Leu Val Lys Asn Ala Gly Leu Glu
450 455 460

Ser Ala Val Ile Ile Asp Tyr Leu Ile Lys Gln Asn Asn Lys Glu Leu
465 470 475 480

Ile Tyr Asn Val Glu Ala Met Ser Tyr Ala Asn Ala Phe Ala Ala Gly
485 490 495

Val Ile Asp Pro Ala Lys Val Val Arg Ile Ala Phe Glu Thr Ala Ile
500 505 510

Ser Val Ala Ser Val Leu Ile Thr Thr Glu Ser Met Ile Val Asp Ile
515 520 525

Pro Asn Lys Asp Glu Asn Ala Ser Ser Pro Met Gly Ala Gly Gly Met
530 535 540

Gly Arg Met Asn Asp Phe
545 550

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 568 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Leu Arg Leu Ala Arg Lys Gly Leu Gln Thr Ala Val Val Arg Ser
1 5 10 15

Tyr Ala Lys Asp Val Lys Phe Gly Ala Glu Gly Arg Gln Ala Met Leu
20 25 30

Val Gly Val Asn Leu Leu Ala Asp Ala Val Ser Val Thr Met Gly Pro
35 40 45

Lys Gly Arg Asn Val Ile Ile Glu Gln Ser Trp Gly Ser Pro Lys Ile
50 55 60

Thr Lys Asp Gly Val Thr Val Ala Lys Ser Ile Asp Leu Lys Asp Lys
65 70 75 80

Tyr Gln Asn Leu Gly Ala Lys Leu Ile Gln Asp Val Ala Asn Lys Ala
85 90 95

```
Asn Glu Glu Ala Gly Asp Gly Thr Thr Cys Ala Thr Val Leu Thr Arg
            100                 105                 110

Ala Ile Ala Lys Glu Gly Phe Glu Arg His Ser Ser Arg Gly Asn Ala
        115                 120                 125

Val Glu Ile Arg Arg Gly Val Met Asn Ala Val Glu Val Val Val Ala
    130                 135                 140

Glu Leu Lys Lys Ile Ser Lys Lys Val Thr Thr Pro Glu Glu Ile Ala
145                 150                 155                 160

Gln Val Ala Thr Ile Ser Ala Asn Gly Asp Thr Val Val Gly Asn Leu
                165                 170                 175

Ile Ser Asp Ala Met Lys Lys Val Gly Thr Thr Gly Val Ile Thr Val
            180                 185                 190

Lys Asp Gly Lys Thr Leu Asn Asp Gln Leu Glu Leu Ile Glu Gly Met
        195                 200                 205

Lys Phe Asp Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Thr Ser Ala Lys
    210                 215                 220

Gly Ala Lys Val Glu Tyr Glu Lys Ala Leu Val Leu Leu Ser Glu Lys
225                 230                 235                 240

Lys Ile Ser Gln Val Gln Asp Ile Val Pro Ala Leu Glu Leu Ala Asn
                245                 250                 255

Lys Leu Arg Arg Pro Leu Val Ile Ile Ala Glu Asp Val Asp Gly Glu
            260                 265                 270

Ala Leu Thr Thr Leu Val Leu Asn Arg Leu Lys Val Gly Leu Gln Val
        275                 280                 285

Val Ala Ile Lys Ala Pro Gly Phe Gly Asp Asn Arg Lys Asn Ala Leu
    290                 295                 300

Lys Asp Met Gly Ile Ala Thr Gly Ala Ser Ile Phe Gly Asp Glu Thr
305                 310                 315                 320

Leu Asp Leu Arg Leu Glu Asp Ile Thr Ala Asn Asp Leu Gly Glu Val
                325                 330                 335

Asp Glu Val Thr Ile Thr Lys Asp Asp Thr Leu Leu Leu Arg Gly Arg
            340                 345                 350

Gly Asp Gln Thr Glu Ile Glu Lys Arg Ile Glu Glu Ile Thr Asp Glu
        355                 360                 365

Ile Glu Arg Ser Thr Ser Asp Tyr Glu Lys Glu Lys Leu Asn Glu Arg
    370                 375                 380

Leu Ala Lys Leu Ser Lys Gly Val Ala Val Leu Lys Ile Gly Gly Gly
```

```
385                 390                 395                 400

Ser Glu Val Glu Val Gly Glu Lys Lys Asp Arg Val Thr Asp Ala Leu
                405                 410                 415

Cys Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly
            420                 425                 430

Val Ala Leu Leu Arg Ser Leu Thr Ala Leu Lys Asn Tyr Lys Ala Ala
        435                 440                 445

Asn Glu Asp Gln Gln Ile Gly Val Asn Ile Val Lys Lys Ala Leu Thr
    450                 455                 460

Gln Pro Ile Ala Thr Ile Val Lys Asn Ala Gly Leu Glu Pro Ser Ser
465                 470                 475                 480

Ile Ile Asp Glu Val Thr Gly Asn Ser Asn Thr Ser Tyr Gly Tyr Asp
                485                 490                 495

Ala Leu Asn Gly Lys Phe Val Asp Met Phe Glu Ala Gly Ile Ile Asp
            500                 505                 510

Pro Thr Lys Val Val Arg Thr Ala Leu Gln Asp Ala Ser Gly Val Ala
        515                 520                 525

Ser Leu Leu Ala Thr Thr Glu Cys Val Val Thr Glu Ile Pro Lys Glu
    530                 535                 540

Glu Ala Val Gly Gly Pro Ala Gly Gly Met Gly Gly Met Gly Gly Met
545                 550                 555                 560

Gly Gly Met Gly Gly Met Gly Phe
                565
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 576 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Met Phe Arg Leu Pro Val Ser Leu Ala Arg Ser Ser Ile Ser Arg Gln
1               5                   10                  15

Leu Ala Met Arg Gly Tyr Ala Lys Asp Val Arg Phe Gly Pro Glu Val
            20                  25                  30
```

```
Arg Ala Met Met Leu Gln Gly Val Asp Val Leu Ala Asp Ala Val Ala
        35                  40                  45

Val Thr Met Gly Pro Lys Gly Arg Asn Val Ile Ile Glu Gln Ser Val
    50                  55                  60

Gly Leu Ala Lys Ile Thr Lys Asp Gly Val Thr Val Ala Lys Ser Ile
65                  70                  75                  80

Glu Leu Lys Asp Lys Phe Gln Asn Ile Gly Ala Lys Leu Val Gln Asp
                85                  90                  95

Leu Ala Asn Asn Thr Asn Glu Glu Ala Gly Asp Gly Thr Thr Thr Ala
            100                 105                 110

Thr Phe Leu Ala Arg Ala Ile Ala Lys Glu Gly Phe Glu Lys Ile Ser
        115                 120                 125

Lys Gly Gly Asn Pro Val Glu Ile Arg Arg Gly Val Met Leu Ala Val
    130                 135                 140

Glu Thr Val Lys Asp Asn Leu Lys Thr Met Ser Arg Pro Val Ser Thr
145                 150                 155                 160

Pro Glu Glu Ile Ala Gln Val Ala Thr Ile Ser Ala Asn Gly Asp Arg
                165                 170                 175

Glu Ile Gly Asn Gly Lys Val Ser Val Ser Glu Ala Met Lys Lys Val
            180                 185                 190

Gly Arg Asp Gly Val Ile Thr Val Lys Asp Gly Lys Thr Leu Thr Asp
        195                 200                 205

Glu Leu Glu Val Ile Glu Gly Thr Met Arg Phe Asp Arg Gly Tyr Ile
    210                 215                 220

Ser Pro Tyr Phe Ile Asn Ser Ser Lys Gly Ala Lys Val Glu Phe Gln
225                 230                 235                 240

Asp Ala Leu Leu Leu Leu Ser Glu Lys Lys Ile Ser Ser Val Ala Glu
                245                 250                 255

His His Ser Pro Leu Trp Arg Leu Ala Ser Arg Arg Thr Arg Lys Pro
            260                 265                 270

Leu Val Ile Ile Ala Glu Asp Ile Asp Gly Glu Ala Leu Ser Thr Leu
        275                 280                 285

Val Val Asn Arg Leu Lys Ile Gly Leu Gln Val Ala Ala Val Lys Ala
    290                 295                 300

Pro Gly Phe Gly Asp Asn Arg Lys Ser Thr Leu Thr Asp Met Ala Thr
305                 310                 315                 320

Ser Gly Gly Ile Val Phe Gly Asp Asp Val Ser Leu Val Lys Leu Glu
```

```
                325                 330                 335

Asp Val Lys Val Ser Asp Leu Gly Gln Val Gly Glu Val Val Ile Thr
            340                 345                 350

Lys Asp Asp Thr Leu Leu Leu Lys Gly Lys Gly Lys Lys Asp Asp Val
        355                 360                 365

Leu Arg Arg Ala Asn Gln Ile Arg Thr Lys Ile Glu Asp Thr Thr Ser
    370                 375                 380

Glu Tyr Glu Lys Glu Lys Leu Gln Glu Arg Leu Ala Arg Leu Ala Ser
385                 390                 395                 400

Gly Val Ala Leu Arg Val Gly Gly Ser Ser Glu Val Glu Val Asn Glu
                405                 410                 415

Lys Lys Asp Arg Val His Asp Ala Leu Asn Ala Thr Arg Ala Ala Val
            420                 425                 430

Glu Glu Gly Ile Val Pro Gly Gly Gly Arg Pro Leu Leu Arg Cys Ile
        435                 440                 445

Glu Lys Leu Glu Gly Val Glu Thr Thr Asn Glu Asp Gln Lys Leu Gly
    450                 455                 460

Val Glu Ile Val Arg Arg Ala Leu Arg Met Pro Cys Met Thr Ile Ala
465                 470                 475                 480

Lys Asn Ala Gly Val Asp Gly Ala Met Val Val Ala Lys Val Glu Asn
                485                 490                 495

Gln Ala Gly Asp Tyr Gly Tyr Asp Ala Lys Gly Glu Tyr Gly Asn Leu
            500                 505                 510

Ile Glu Lys Gly Ile Ile Asp Pro Thr Lys Val Val Arg Thr Ala Ile
        515                 520                 525

Thr Asp Ala Ser Gly Val Ala Ser Leu Leu Thr Thr Ala Glu Ala Val
    530                 535                 540

Val Thr Glu Ile Pro Lys Glu Asp Gly Ala Pro Ala Met Pro Gly Met
545                 550                 555                 560

Gly Gly Met Gly Gly Met Gly Gly Met Gly Gly Met Gly Gly Met Met
                565                 570                 575
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 573 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Met Leu Arg Leu Pro Thr Val Phe Arg Gln Met Arg Pro Val Ser Arg
1               5                  10                  15

Val Leu Ala Pro His Leu Thr Arg Ala Tyr Ala Lys Asp Val Lys Phe
            20                  25                  30

Gly Ala Asp Ala Arg Ala Leu Met Leu Gln Gly Val Asp Leu Leu Ala
        35                  40                  45

Asp Ala Val Ala Val Thr Met Gly Pro Lys Gly Arg Thr Val Ile Ile
    50                  55                  60

Glu Gln Gly Trp Gly Ser Pro Lys Val Thr Lys Asp Gly Val Thr Val
65                  70                  75                  80

Ala Lys Ser Ile Asp Leu Lys Asp Lys Tyr Lys Asn Ile Gly Ala Lys
                85                  90                  95

Leu Val Gln Asp Val Ala Asn Asn Thr Asn Glu Glu Ala Gly Asp Gly
            100                 105                 110

Thr Thr Thr Ala Thr Val Leu Ala Arg Ser Ile Ala Lys Glu Gly Phe
        115                 120                 125

Glu Lys Ile Ser Lys Gly Ala Asn Pro Val Glu Ile Arg Arg Gly Val
    130                 135                 140

Met Leu Ala Val Asp Ala Val Ile Ala Glu Leu Lys Lys Gln Ser Lys
145                 150                 155                 160

Pro Val Thr Thr Pro Glu Glu Ile Ala Gln Val Ala Thr Ile Ser Ala
                165                 170                 175

Asn Gly Asp Lys Glu Ile Gly Asn Ile Ile Ser Asp Ala Met Lys Lys
            180                 185                 190

Val Gly Arg Lys Gly Val Ile Thr Val Lys Asp Gly Lys Thr Leu Asn
        195                 200                 205

Asp Glu Leu Glu Ile Ile Glu Gly Met Lys Phe Asp Arg Gly Tyr Ile
    210                 215                 220

Ser Pro Tyr Phe Ile Asn Thr Ser Lys Gly Gln Lys Cys Glu Phe Gln
225                 230                 235                 240

Asp Ala Tyr Val Leu Leu Ser Glu Lys Lys Ile Ser Ser Ile Gln Ser
                245                 250                 255
```

```
Ile Val Pro Ala Leu Glu Ile Ala Asn Ala His Arg Lys Pro Leu Val
            260                 265                 270

Ile Ile Ala Glu Asp Val Asp Gly Glu Ala Leu Ser Thr Leu Val Leu
        275                 280                 285

Asn Arg Leu Lys Val Gly Leu Gln Val Val Ala Val Lys Ala Pro Gly
    290                 295                 300

Phe Gly Asp Asn Arg Lys Asn Gln Leu Lys Asp Met Ala Ile Ala Thr
305                 310                 315                 320

Gly Gly Ala Val Phe Gly Glu Glu Gly Leu Thr Leu Asn Leu Glu Asp
                325                 330                 335

Val Gln Pro His Asp Leu Gly Lys Val Gly Glu Val Ile Val Thr Lys
            340                 345                 350

Asp Asp Ala Met Leu Leu Lys Gly Lys Gly Asp Lys Ala Gln Ile Glu
        355                 360                 365

Lys Arg Ile Gln Glu Ile Ile Glu Gln Leu Asp Val Thr Thr Ser Glu
    370                 375                 380

Tyr Glu Lys Glu Lys Leu Asn Glu Arg Leu Ala Lys Leu Ser Asp Gly
385                 390                 395                 400

Val Ala Val Leu Lys Val Gly Gly Thr Ser Asp Val Glu Val Asn Glu
                405                 410                 415

Lys Lys Asp Arg Val Thr Asp Ala Leu Asn Ala Thr Arg Ala Ala Val
            420                 425                 430

Glu Glu Gly Ile Val Leu Gly Gly Gly Cys Ala Leu Leu Arg Cys Ile
        435                 440                 445

Pro Ala Leu Asp Ser Leu Thr Pro Ala Asn Glu Asp Gln Lys Ile Gly
    450                 455                 460

Ile Glu Ile Ile Lys Arg Thr Leu Lys Ile Pro Ala Met Thr Ile Ala
465                 470                 475                 480

Lys Asn Ala Gly Val Glu Gly Ser Leu Ile Val Glu Lys Ile Met Gln
                485                 490                 495

Ser Ser Ser Glu Val Gly Tyr Asp Ala Met Ala Gly Asp Phe Val Asn
            500                 505                 510

Met Val Glu Lys Gly Ile Ile Asp Pro Thr Lys Val Val Arg Thr Ala
        515                 520                 525

Leu Leu Asp Ala Ala Gly Val Ala Ser Leu Leu Thr Thr Ala Glu Val
    530                 535                 540

Val Val Thr Glu Ile Pro Lys Glu Glu Lys Asp Pro Gly Met Gly Ala
```

545 550 555 560

Met Gly Gly Met Gly Gly Gly Met Gly Gly Gly Met Phe
565 570

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 577 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Tyr Arg Phe Ala Ser Asn Leu Ala Ser Lys Ala Arg Ile Ala Gln
1 5 10 15

Asn Ala Arg Gln Val Ser Ser Arg Met Ser Trp Ser Arg Asn Tyr Ala
20 25 30

Ala Lys Glu Ile Lys Phe Gly Val Glu Ala Arg Ala Leu Met Leu Lys
35 40 45

Gly Val Glu Asp Leu Ala Asp Ala Val Lys Val Thr Met Gly Pro Lys
50 55 60

Gly Arg Asn Val Val Ile Glu Gln Ser Trp Gly Ala Pro Lys Val Thr
65 70 75 80

Lys Asp Gly Val Thr Val Ala Lys Ser Ile Glu Phe Lys Asp Lys Ile
85 90 95

Lys Asn Val Gly Ala Ser Leu Val Lys Gln Val Ala Asn Ala Thr Asn
100 105 110

Asp Val Ala Gly Asp Gly Thr Thr Cys Ala Thr Val Leu Thr Arg Ala
115 120 125

Ile Phe Ala Glu Gly Cys Lys Ser Val Ala Ala Gly Met Asn Ala Met
130 135 140

Asp Leu Arg Arg Gly Ile Ser Met Ala Val Asp Ala Val Val Thr Asn
145 150 155 160

Leu Lys Ser Lys Ala Arg Met Ile Ser Thr Ser Glu Glu Ile Ala Gln
165 170 175

Val Gly Thr Ile Ser Ala Asn Gly Glu Arg Glu Ile Gly Glu Leu Ile
180 185 190

```
Ala Lys Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr Ile Gln
        195                 200                 205

Asp Gly Lys Thr Leu Phe Asn Glu Leu Glu Val Val Glu Gly Met Lys
    210                 215                 220

Leu Asp Arg Gly Tyr Thr Ser Pro Tyr Phe Ile Thr Asn Gln Lys Thr
225                 230                 235                 240

Gln Lys Cys Glu Leu Asp Asp Pro Leu Ile Leu Ile His Glu Lys Lys
                245                 250                 255

Ile Ser Ser Ile Asn Ser Ile Val Lys Val Leu Glu Leu Ala Leu Lys
            260                 265                 270

Arg Gln Arg Pro Leu Leu Ile Val Ser Glu Asp Val Glu Ser Asp Ala
        275                 280                 285

Leu Ala Thr Leu Ile Leu Asn Lys Leu Arg Ala Gly Ile Lys Val Cys
    290                 295                 300

Ala Ile Lys Ala Pro Gly Phe Gly Glu Asn Arg Lys Ala Asn Leu Gln
305                 310                 315                 320

Asp Leu Ala Ala Leu Thr Gly Gly Glu Val Ile Thr Asp Glu Leu Gly
                325                 330                 335

Met Asn Leu Glu Lys Val Asp Leu Ser Met Leu Gly Thr Cys Lys Lys
            340                 345                 350

Val Thr Val Ser Lys Asp Asp Thr Val Ile Leu Asp Gly Ala Gly Asp
        355                 360                 365

Lys Lys Gly Ile Glu Glu Arg Cys Glu Gln Ile Arg Ser Ala Ile Glu
    370                 375                 380

Leu Ser Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu Arg Leu Ala
385                 390                 395                 400

Lys Leu Ser Gly Gly Val Ala Val Leu Lys Ile Gly Gly Ala Ser Glu
                405                 410                 415

Ala Glu Val Gly Glu Lys Lys Asp Arg Val Thr Asp Ala Leu Asn Ala
            420                 425                 430

Thr Lys Ala Ala Val Glu Glu Gly Ile Leu Pro Gly Gly Gly Val Ala
        435                 440                 445

Leu Leu Tyr Ala Ala Arg Glu Leu Glu Lys Leu Pro Thr Ala Asn Phe
    450                 455                 460

Asp Gln Lys Ile Gly Val Gln Ile Ile Gln Asn Ala Leu Lys Thr Pro
465                 470                 475                 480

Val Tyr Thr Ile Ala Ser Asn Ala Gly Val Glu Gly Ala Val Ile Val
```

```
                485                 490                 495

Gly Lys Leu Leu Glu Gln Asp Asn Pro Asp Leu Gly Tyr Asp Ala Ala
            500                 505                 510

Lys Gly Glu Tyr Val Asp Met Val Lys Ala Gly Ile Ile Asp Pro Leu
        515                 520                 525

Lys Val Ile Arg Thr Ala Leu Val Asp Ala Ala Ser Val Ser Ser Leu
    530                 535                 540

Leu Thr Thr Thr Glu Ala Val Val Val Asp Leu Pro Lys Asp Glu Ser
545                 550                 555                 560

Glu Ser Gly Ala Ala Gly Gly Gly Met Gly Gly Met Val Val Met Asp
                565                 570                 575

Tyr
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 576 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Met Tyr Arg Ala Ala Ala Ser Leu Ala Ser Lys Ala Arg Gln Ala Gly
1               5                   10                  15

Ser Ser Ser Ala Ala Arg Gln Val Gly Ser Arg Leu Ala Trp Ser Arg
            20                  25                  30

Asn Tyr Ala Ala Lys Asp Ile Lys Phe Gly Val Glu Ala Arg Ala Leu
        35                  40                  45

Met Leu Arg Gly Val Glu Glu Leu Ala Asp Ala Val Lys Val Thr Met
    50                  55                  60

Gly Pro Lys Gly Arg Asn Val Val Ile Glu Gln Ser Phe Gly Ala Pro
65                  70                  75                  80

Lys Val Thr Lys Asp Gly Val Thr Val Ala Lys Ser Ile Glu Phe Lys
                85                  90                  95

Asp Arg Val Lys Asn Val Gly Ala Ser Leu Val Lys Gln Val Ala Asn
            100                 105                 110
```

```
Ala Thr Asn Asp Asn Ala Gly Asp Gly Thr Thr Cys Ala Thr Val Leu
        115                 120                 125

Thr Lys Ala Ile Phe Thr Glu Gly Cys Lys Ser Val Ala Ala Gly Met
    130                 135                 140

Asn Ala Met Asp Leu Arg Arg Gly Ile Ser Met Ala Val Asp Ala Val
145                 150                 155                 160

Val Thr Asn Leu Lys Gly Met Ala Arg Met Ile Ser Thr Ser Glu Glu
                165                 170                 175

Ile Ala Gln Val Gly Thr Ile Ser Ala Asn Gly Glu Arg Glu Ile Gly
            180                 185                 190

Glu Leu Ile Ala Lys Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile
        195                 200                 205

Thr Ile Ala Asp Gly Asn Thr Leu Tyr Asn Glu Leu Glu Val Val Glu
    210                 215                 220

Gly Met Lys Leu Asp Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Thr Asn
225                 230                 235                 240

Ser Lys Ala Gln Lys Cys Glu Pro Glu Asp Pro Leu Ile Leu Ile His
                245                 250                 255

Asp Arg Lys Val Thr Asn Met His Ala Val Val Lys Val Leu Glu Met
            260                 265                 270

Ala Leu Lys Lys Gln Arg Pro Leu Leu Ile Val Ala Glu Asp Val Glu
        275                 280                 285

Ser Glu Ala Leu Gly Thr Leu Ile Ile Asn Lys Leu Arg Ala Gly Ile
    290                 295                 300

Lys Val Cys Ala Val Lys Ala Pro Gly Phe Gly Glu Asn Arg Lys Ala
305                 310                 315                 320

Asn Leu Gln Asp Leu Ala Ile Leu Thr Gly Gly Glu Val Ile Thr Glu
                325                 330                 335

Glu Leu Gly Met Asn Leu Glu Asn Val Glu Pro His Met Leu Gly Ser
            340                 345                 350

Cys Lys Lys Val Thr Val Ser Lys Asp Asp Thr Val Ile Leu Asp Gly
        355                 360                 365

Ala Gly Asp Lys Lys Ser Ile Glu Glu Arg Ala Asp Gln Ile Arg Ser
    370                 375                 380

Ala Val Glu Asn Ser Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu
385                 390                 395                 400

Arg Leu Ala Lys Leu Ser Gly Gly Val Ala Val Leu Lys Ile Gly Gly
```

405 410 415

Ala Ser Glu Ala Glu Val Gly Glu Lys Lys Asp Arg Val Thr Asp Ala
420 425 430

Leu Asn Ala Thr Lys Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly
435 440 445

Gly Val Ala Leu Leu Tyr Ala Ser Lys Glu Leu Asp Lys Leu Gln Thr
450 455 460

Ala Asn Phe Asp Gln Lys Ile Gly Val Gln Ile Ile Gln Asn Ala Leu
465 470 475 480

Lys Thr Pro Val His Thr Ile Ala Ser Asn Ala Gly Val Glu Gly Ala
485 490 495

Val Val Val Gly Lys Leu Leu Glu Gln Gly Asn Thr Asp Leu Gly Tyr
500 505 510

Asp Ala Ala Lys Asp Glu Tyr Val Asp Met Val Lys Ala Gly Ile Ile
515 520 525

Asp Pro Leu Lys Val Ile Arg Thr Ala Leu Val Asp Ala Ala Ser Val
530 535 540

Ser Ser Leu Met Thr Thr Thr Glu Ser Ile Ile Val Glu Ile Pro Lys
545 550 555 560

Glu Glu Ala Pro Ala Pro Ala Met Gly Gly Met Gly Gly Met Asp Tyr
565 570 575

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 587 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met Ala Ser Thr Asn Ala Leu Ser Ser Thr Ser Ile Leu Arg Ser Pro
1 5 10 15

Thr Asn Gln Ala Gln Thr Ser Leu Ser Lys Lys Val Lys Gln His Gly
20 25 30

Arg Val Asn Phe Arg Gln Lys Pro Asn Arg Phe Val Val Lys Ala Ala
35 40 45

```
Ala Lys Asp Ile Ala Phe Asp Gln His Ser Arg Ser Ala Met Gln Ala
    50                  55                  60

Gly Ile Asp Lys Leu Ala Asp Ala Val Gly Leu Thr Leu Gly Pro Arg
65                  70                  75                  80

Gly Arg Asn Val Val Leu Asp Glu Phe Gly Ser Pro Lys Val Val Asn
                85                  90                  95

Asp Gly Val Thr Ile Ala Arg Ala Ile Glu Leu Pro Asp Pro Met Glu
            100                 105                 110

Asn Ala Gly Ala Ala Leu Ile Arg Glu Val Ala Ser Lys Thr Asn Asp
        115                 120                 125

Ser Ala Gly Asp Gly Thr Thr Thr Ala Ser Ile Leu Ala Arg Glu Ile
    130                 135                 140

Ile Lys Leu Gly Leu Leu Asn Val Thr Ser Gly Ala Asn Pro Val Ser
145                 150                 155                 160

Ile Lys Lys Gly Ile Asp Lys Thr Val Ala Ala Leu Val Glu Glu Leu
                165                 170                 175

Glu Lys Leu Ala Arg Pro Val Lys Gly Gly Asp Asp Ile Lys Ala Val
            180                 185                 190

Ala Thr Ile Ser Ala Gly Asn Asp Glu Leu Ile Gly Lys Met Ile Ala
        195                 200                 205

Glu Ala Ile Asp Lys Val Gly Pro Asp Gly Val Leu Ser Ile Glu Ser
    210                 215                 220

Ser Asn Ser Phe Glu Thr Thr Val Glu Val Glu Glu Gly Met Glu Ile
225                 230                 235                 240

Asp Arg Gly Tyr Ile Ser Pro Gln Phe Val Thr Asn Pro Glu Lys Ser
                245                 250                 255

Ile Val Glu Phe Glu Asn Ala Arg Val Leu Ile Thr Asp Gln Lys Ile
            260                 265                 270

Ser Ala Ile Lys Asp Ile Ile Pro Leu Leu Glu Lys Thr Thr Gln Leu
        275                 280                 285

Arg Ala Pro Leu Leu Ile Ile Ser Glu Asp Ile Thr Gly Glu Ala Leu
    290                 295                 300

Ala Thr Leu Val Val Asn Lys Leu Arg Gly Ile Leu Asn Val Ala Ala
305                 310                 315                 320

Ile Lys Ala Pro Gly Phe Gly Glu Arg Arg Lys Ala Leu Leu Gln Asp
                325                 330                 335
```

```
Ile Ala Ile Leu Thr Gly Ala Glu Phe Gln Ala Ser Asp Leu Gly Leu
            340                 345                 350

Leu Val Glu Asn Thr Thr Ile Glu Gln Leu Gly Leu Ala Arg Lys Val
        355                 360                 365

Thr Ile Ser Lys Asp Ser Thr Thr Ile Ile Ala Asp Ala Ala Ser Lys
    370                 375                 380

Asp Glu Leu Gln Ser Arg Val Ala Gln Leu Lys Lys Glu Leu Ser Glu
385                 390                 395                 400

Thr Asp Ser Ile Tyr Asp Ser Glu Lys Leu Ala Glu Arg Ile Ala Lys
                405                 410                 415

Leu Ser Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu Thr
            420                 425                 430

Glu Leu Glu Asp Arg Lys Leu Arg Ile Glu Asp Ala Lys Asn Ala Thr
        435                 440                 445

Phe Ala Ala Ile Glu Glu Gly Ile Val Pro Gly Gly Gly Thr Ala Leu
    450                 455                 460

Val His Leu Ser Gly Tyr Val Pro Ala Ile Lys Glu Lys Leu Glu Asp
465                 470                 475                 480

Ala Asp Glu Arg Leu Gly Ala Asp Ile Val Gln Lys Ala Leu Val Ala
                485                 490                 495

Pro Ala Ala Leu Ile Ala Gln Asn Ala Gly Ile Glu Gly Glu Val Val
            500                 505                 510

Val Glu Lys Ile Lys Asn Gly Glu Trp Glu Val Gly Tyr Asn Ala Met
        515                 520                 525

Thr Asp Thr Tyr Glu Asn Leu Val Glu Ser Gly Val Ile Asp Pro Ala
    530                 535                 540

Lys Val Thr Arg Cys Ala Leu Gln Asn Ala Ala Ser Val Ala Gly Met
545                 550                 555                 560

Val Leu Thr Thr Gln Ala Ile Val Val Glu Lys Pro Lys Pro Lys Ala
                565                 570                 575

Ala Val Ala Ala Ala Pro Gln Gly Leu Thr Ile
            580                 585
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 545 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Met Ala Lys Asp Ile Lys Phe Gly Glu Glu Ala Arg Arg Ala Met Leu
1               5                  10                  15

Arg Gly Val Asn Ala Leu Ala Asp Ala Val Lys Val Thr Leu Gly Pro
            20                  25                  30

Lys Gly Arg Asn Val Val Leu Glu Lys Ser Phe Gly Ala Pro Thr Ile
        35                  40                  45

Thr Lys Asp Gly Val Thr Val Ala Lys Glu Ile Glu Leu Glu Asp Lys
    50                  55                  60

Phe Glu Asn Met Gly Ala Gln Leu Val Lys Glu Val Ala Ser Lys Thr
65                  70                  75                  80

Asn Asp Val Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Gln
                85                  90                  95

Ala Ile Val Lys Glu Gly Leu Lys Asn Val Ala Ala Gly Ala Asn Pro
            100                 105                 110

Met Asp Leu Arg Arg Gly Ile Asp Lys Ala Val Asp Ala Val Val Glu
        115                 120                 125

Glu Leu Lys Ala Ile Ala Lys Pro Val Glu Thr Lys Glu Glu Ile Ala
    130                 135                 140

Gln Val Ala Thr Ile Ser Ala Asn Gly Asp Glu Glu Ile Gly Glu Leu
145                 150                 155                 160

Ile Ala Glu Ala Met Glu Lys Val Gly Lys Glu Gly Val Ile Thr Val
                165                 170                 175

Glu Glu Gly Lys Thr Leu Glu Thr Glu Leu Glu Val Val Glu Gly Met
            180                 185                 190

Gln Phe Asp Arg Gly Tyr Ile Ser Pro Tyr Phe Ile Thr Asp Ser Glu
        195                 200                 205

Lys Gln Lys Ala Glu Leu Glu Asp Pro Leu Ile Leu Leu Thr Asp Lys
    210                 215                 220

Lys Ile Ser Asn Ile Gln Asp Leu Leu Pro Val Leu Glu Glu Val Ala
225                 230                 235                 240

Gln Ala Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly Glu
                245                 250                 255
```

```
Ala Leu Ala Thr Leu Val Val Asn Lys Leu Arg Gly Thr Leu Lys Val
            260                 265                 270

Val Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Ala Met Leu
        275                 280                 285

Gln Asp Ile Ala Ile Leu Thr Gly Gly Gln Val Ile Ser Glu Glu Leu
    290                 295                 300

Gly Leu Ser Leu Glu Asp Ala Thr Leu Glu Asp Leu Gly Gln Ala Lys
305                 310                 315                 320

Lys Val Val Val Thr Lys Asp Asp Thr Thr Ile Val Asp Gly Ala Gly
                325                 330                 335

Asp Ala Ala Ile Ala Gly Arg Val Ala Gln Ile Arg Ser Gln Ile Glu
            340                 345                 350

Glu Ser Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ala Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Thr Glu
    370                 375                 380

Val Glu Leu Lys Glu Arg Lys Asp Arg Val Glu Asp Ala Leu Asn Ala
385                 390                 395                 400

Thr Arg Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly Val Ala
                405                 410                 415

Leu Leu Arg Ala Ala Pro Ala Leu Asp Lys Leu Lys Thr Glu Asn Gly
            420                 425                 430

Asp Glu Ala Thr Gly Val Asn Ile Val Leu Arg Ala Leu Glu Ala Pro
        435                 440                 445

Leu Arg Gln Ile Ala Glu Asn Ala Gly Leu Glu Gly Ser Val Val Val
    450                 455                 460

Glu Lys Val Lys Asn Ser Glu Ala Gly Gly Tyr Asn Ala Ala Thr Gly
465                 470                 475                 480

Glu Tyr Val Asp Met Ile Ala Ala Gly Ile Ile Asp Pro Thr Lys Val
                485                 490                 495

Thr Arg Ser Ala Leu Gln Asn Ala Ala Ser Val Ala Ser Leu Met Leu
            500                 505                 510

Thr Thr Glu Ala Val Val Val Asp Lys Pro Glu Lys Glu Ala Ala Pro
        515                 520                 525

Ala Gly Met Pro Gly Met Met Gly Gly Met Gly Gly Met Gly Gly Met
    530                 535                 540

Met
```

545

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CATATGGCNG CNAAAGAYGT AAAA 24

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TGATCACATC ATNCCNCCCA TNCC 24

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CATATGGCAA AAGAAATHAA RTTY 24

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGATCANCCN CCCATNCCNC CCAT 24

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTAAAACGAC GGCCAG 16

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CAGGAAACAG CTATGAC 17

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCAACCATCA CGAAAGA 17

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ACGGGTCACT TTGGTTG 17

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TTACTAATGA CGGGGTA 17

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TTACCAATGA CGGTGTG 17

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ACAGGGTCAA TGATTCC 17

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ACTGGATCAA TGATACC 17

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CCGTACCGTG CTCTGAC 17

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ACCACGTTTC AGATCCA 17

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GACAGTTTCG CGGCAAC 17

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CTCAGAACGA AGATCAG 17

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GGTATGCAGT TCGACCG 17

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CCGTGTTGGT CAAATCC 17

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGTAACTACG GTTACAA 17

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GAGGCCACTT CTTTCAC 17

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GGCTTCCAGC ACTGGCA 17

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AACTTCAGTC GCAGCAC 17

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CCTTGAAAGC CATTGCT 17

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GCTACACGTG CAGCCGT 17

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GCTGCAACAG GTGAGTG 17

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TCATGAACAA TGGCTTG 17

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

ACGAAGCACA ATGTTAC 17

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

ATCACTAAAG ATGGTGT 17

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GCAGTTGCCG CAGCAGT 17

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GCTACTCGTG CAGCTGT 17

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GTTCTCCGTG CTTTGGA 17

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GCACCTGCTG TGACGTT 17

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

TCTTCGATGG TGATGAC 17

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GGCAAGAGCT GTTCCGC 17

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CTGAGCCAGT ACGGTTG 17

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GTACTGCAGA GCGGAAC 17

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ACCGTCTTCA ACGGTGA 17

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GTTATCATTG CTGAAGA 17

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ACGGTACCGC CGGTCAG 17

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGGGCCAGG CTAAACG 17

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CGACTGAAGT TGAAATG 17

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GCTGTTGAAG AACTGAA 17

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GTCTTCAACG GTGATCA 17

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

TCTTCTACCG CAGCACG 17

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

CTCTTGATTA TTGCGGA 17

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TTGTTCAAAA CAAGAGT 17

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

CGATTATTGT AGAAGGT 17

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

CTTGATAACC GCAACAC 17

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

TCCAAAGCAC GGAGAAC 17

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

GTGTCAAACA TCCAAGA 17

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TCTTCGATGG TAATCAC 17

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

GCAATAATGA GTAATGG 17

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

ACAGTAATTG TTGAAGG 17

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

CAGTGCAATA CGGTTAG 17

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

AGCTTCCAGA ACCGGCA 17

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CTGATCATCG CTGAAGA 17

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

ACGGTTATTG TAGAAG 16

INTERNATIONAL SEARCH REPORT

Inte. .onal Application No
PCT/CA 98/01203

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/31 C07K14/315 C07K19/00 C12N15/70 C12N1/21
A61K39/09

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LUTHER E. LINDLER ET AL.: "Nucleotide sequence of the Salmonella typhi groEL heat shock gene" MICROBIAL PATHOGENESIS, vol. 17, no. 4, October 1994, pages 271-275, XP002099747 see the whole document --- -/--	1-31

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 April 1999	27/04/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	MONTERO LOPEZ, B

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inte. .onal Application No
PCT/CA 98/01203

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	HAMEL J ET AL: "Heat shock response of Streptococcus pneumoniae: identification of immunoreactive stress proteins." MICROBIAL PATHOGENESIS, (1997 JUL) 23 (1) 11-21. JOURNAL CODE: MIC. ISSN: 0882-4010., XP002099748 ENGLAND: United Kingdom see page 12, right-hand column, paragraph 2 - page 13, right-hand column, paragraph 1 see page 16, right-hand column, paragraph 1 - page 18, left-hand column, paragraph 2 ---	9,12, 25-29
X	BENKIRANE R ET AL: "Identification of a Streptococcus suis 60-kDa heat - shock protein using western blotting." FEMS MICROBIOLOGY LETTERS, (1997 AUG 15) 153 (2) 379-85. JOURNAL CODE: FML. ISSN: 0378-1097., XP002099749 Netherlands see page 381, right-hand column, paragraph 2 - page 384, right-hand column, paragraph 1 ---	9,10
A	WO 96 40928 A (IAF BIOVAC INC.) 19 December 1996 see page 6, line 35 - page 8, line 16 see page 15, line 15 - page 33, line 23 ---	1-31
P,X	WO 98 18931 A (HUMAN GENOME SCIENCES, INC.) 7 May 1998 see page 4, line 4 - page 6, line 2 see page 16, line 16 - line 19 see page 16, line 23 - page 18, line 27 see page 21, line 19 - page 29, line 11 see page 37, line 19 - page 41, line 13 see page 70; table 2 see sequence SEQ ID NO:77 ---	1,3-9, 11-31
P,X	Trpro Database Entry 033733 Accession number 033733; 1 January 1998 POHL B. ET AL. XP002099751 see the whole document --- -/--	10,11, 14-17

1

INTERNATIONAL SEARCH REPORT

Inte ional Application No
PCT/CA 98/01203

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	LEMOS J A ET AL: "Expression of heat - shock proteins in Streptococcus pyogenes and their immunoreactivity with sera from patients with streptococcal diseases." JOURNAL OF MEDICAL MICROBIOLOGY, (1998 AUG) 47 (8) 711-5. JOURNAL CODE: J2N. ISSN: 0022-2615., XP002099750 ENGLAND: United Kingdom see page 712, right-hand column, paragraph 3 - page 715, right-hand column, paragraph 1	10,25-28

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA 98/ 01203

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 29 and 30 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte .onal Application No
PCT/CA 98/01203

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9640928 A	19-12-1996	AU 700080 B	17-12-1998
		AU 5682896 A	30-12-1996
		CA 2224015 A	19-12-1996
		CN 1192241 A	02-09-1998
		CZ 9703942 A	15-04-1998
		EP 0832238 A	01-04-1998
		NO 975752 A	06-02-1998
		PL 323781 A	27-04-1998
		SK 168497 A	08-07-1998
WO 9818931 A	07-05-1998	AU 5194598 A	22-05-1998
		AU 6909098 A	22-05-1998
		WO 9818930 A	07-05-1998